

Decision Memo for Lumbar Artificial Disc Replacement (CAG-00292N)

Decision Summary

The Centers for Medicare and Medicaid Services (CMS) has found that lumbar artificial disc replacement (LADR) with the Charite lumbar artificial disc is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, we are issuing a national noncoverage determination for LADR with the Charite lumbar artificial disc for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

[Back to Top](#)

Decision Memo

TO: Administrative File: (CAG-#00292N)
Lumbar Artificial Disc Replacement

FROM:

Steve Phurrough, MD, MPA
Director
Coverage and Analysis Group

Marcel E. Salive, MD, MPH
Director
Division of Medical and Surgical Services

Deirdre O'Connor
Lead Health Policy Analyst, Division of Medical and Surgical Services

Jyme Schafer, MD, MPH
Lead Medical Officer, Division of Medical and Surgical Services

Shamiram Feinglass, MD, MPH
Medical Officer, Division of Items and Devices

SUBJECT: Coverage Decision Memorandum for Lumbar Artificial Disc Replacement

DATE: May 16, 2006

I. Decision

The Centers for Medicare and Medicaid Services (CMS) has found that lumbar artificial disc replacement (LADR) with the Charite lumbar artificial disc is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, we are issuing a national noncoverage determination for LADR with the Charite lumbar artificial disc for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

II. Background

Millions of Americans suffer from pain-related problems (Salovey, Seiber et al. 1992). Low back pain is a common condition, with sixty to eighty percent of U.S. adults afflicted at some time during their life (U.S. Preventive Services Task Force 1996). Low back pain can be defined as symptoms of pain, muscle tension, or stiffness localized below the costal margin and above the inferior gluteal folds, with or without leg pain (Manek, MacGregor 2005). Low back pain can be thought of as being either nonspecific or specific. In specific types of low back pain, the symptoms are caused by pathological conditions such as spinal fractures, cancer, or infection and can be identified and treated appropriately (Manek, MacGregor 2005). Approximately 90% of low back pain is of the nonspecific type (Manek, MacGregor 2005). In nonspecific low back pain, most patient's symptoms resolve satisfactorily within a relatively short time span. In the 5 – 10% of patients whose pain does not satisfactorily resolve, the symptoms can be disabling. Some psychosocial risk factors for the progression to chronicity have been identified (Manek, MacGregor 2005). In general, the social and economic impact of chronic pain is enormous (Salovey, Seiber et al. 1992).

Discovering the cause for nonspecific low back symptoms remains challenging. Haldeman states "...we do not know the origin of low back pain in the majority of cases..." and attributes this conundrum to the unique anatomic complexity of the spine (Haldeman 1999). Neurophysiologic mechanisms of pain sensation are poorly understood, adding to the difficulty in localizing the pain source (Haldeman 1999). Frequently, persistent low back pain is attributed to a damaged intervertebral disc, which bears some of the highest loads in the human body and is almost avascular (Huang, Sandhu 2004). Disc damage, or degeneration, can occur as an ongoing process where ultimately the disc's reparative capacity is overwhelmed, leading to continued changes. Huang and Sandhu stated, "it is not surprising that DDD [degenerative disc disease] is a common phenomenon in middle age and a universal condition in old age." While from a simple mechanical aspect it could be hypothesized that DDD is a cause for pain, disc degeneration is also observed in individuals without pain (Boden, David et al. 1990).

Initial treatment of pain believed to be caused from degenerative disc disease is conservative care. Conservative care can include physical therapy, manipulation, massage, pain medications, and exercise. The majority of patients will have acceptable results with a non-surgical approach. When patients fail conservative care, surgery becomes an option. Until recently in the United States, surgical options available for degenerative disc disease have ranged from discectomies (open or microsurgical) to percutaneous nucleotomies, chemical and thermal nucleolysis and/or spinal fusion (Gibson, Wassell 2005). Spinal fusion has been the predominant surgical treatment for degenerative disc disease (DDD) that does not respond to other treatments. Fusion proposes to relieve pain by eliminating motion in the area of the disc space and/or by disc mechanical load reduction. Nevertheless, the indications for lumbar spinal fusion are variable and not clearly defined (Krismer 2002). These different opinions concerning the indications for back surgery are reflected in the significant regional variation of rates of surgery, surgical techniques used, technical success and rate of fusion (Gibson, Wassell 2005). Satisfactory clinical outcomes can range from 16 to 95% (Gibson, Wassell 2005). Short term relief of pain may perhaps occur with the various types of fusion procedures, but long-term results remain controversial (Bertagnoli, Kumar 2002). Suspected problems include accelerated degeneration of the adjacent lumbar segments, pseudoarthrosis, spinal stenosis and persistent or recurrent low-back pain. In an attempt to overcome these potential long-term problems, the idea of a total artificial disc replacement as a treatment for pain believed secondary to degenerative disc disease has been proposed as an alternative to spinal fusion. As possible added benefits, it has been postulated that total disc replacement may have a protective role on the facet joints, and restore lumbar segment motion (Bertagnoli, Kumar 2002). The artificial disc concept is not new. In the late 1960's, Fernstrom explored the possibility of replacing the intervertebral disc with an artificial disc. Much research and development work has been done since then. The current Charite disc is the third modification of a device first developed in 1982 by Buttner-Janz and Schellnack at the Charite Clinic in the former East Germany. There are other artificial discs in use in other countries and additional disc implants under development. Intervertebral disc replacement design has been problematic due to the three-column structure of the spine, and the three separate joints at each level. The disc is not a true joint, and functions in both mobility and damping, with the center of rotation moving constantly along three axes (Gunzburg, Mayer et al. 2002). Huang and Sandhu suggest the ideal disc replacement would perform the functions of the replaced native disc, which include preservation of physiologic range of motion, transmission of compressive loads across the disc space, protection of the posterior elements (facets) from abnormal loads, and then to function for many years. In general, the current replacement discs that are either approved or under FDA approved trials in the US have metal endplates that affix to the vertebral bony endplates with some mechanism between these two plates that allows for motion in various planes. The Charite disc has an ultra high molecular weight polyethylene insert that sits between the two metal endplates. The metal endplates have external spikes for engagement of the device into the vertebral bony endplates. Components come in various sizes for a close fit to the patient's anatomy. The motion of this disc is thought to be minimally constrained in flexion, extension, lateral bending, and axial rotation. It is constrained in compression. The other disc implants in development in the United States are somewhat similar but can vary in material (metal on polymer or metal on metal), motion design, and method of fixation to vertebral endplate (Santos, Polly et al. 2004). Anderson and Rouleau offered, "The current designs are diverse and, thus far, the effects of their individual characteristics on results are unknown." As of 2004, more than 7,000 patients worldwide had been implanted with the Charite disc (FDA In-depth Statistical Review for Expedited Premarket Approval (PMA) 2004).

The surgical procedure for disc replacement involves an anterior approach for exposure of the spine. With this approach, complications of vessel injury can occur and have the potential to be life threatening (Santos, Polly et al. 2004). On revision surgery, Santos, et al., stated, “Revision surgery for a failed disc arthroplasty is life threatening. Dealing with the scarring around the great vessels is the main challenge. Indeed, the location of vital vascular structures may make it altogether impossible to perform such anterior abdominal exposures.” Other postoperative difficulties such as infection, persistent pain, instability, and osteolysis can occur (Santos, Polly et al. 2004).

The Charite lumbar artificial disc is referred to in numerous ways (SB Charite, Charite SB and sometimes with the number I, II or III) in the many articles reviewed for this analysis. The designation of the number I, II, III denotes the Charite design that was being used at the time. The Charite disc has gone through changes in the design over the years and the numbers designate the design model. The Charite PMA trial used the most current design, the SB Charite III. For the purposes of this document, the term Charite will be used to refer to the implant. If a particular article identifies the design number (I, II, or III), that will be indicated.

III. History of Medicare Coverage

Medicare does not currently have a national coverage determination (NCD) on lumbar artificial disc replacement. Coverage for the procedure is overseen by local Medicare contractors.

Medicare also does not have a NCD for other spinal surgeries for degenerative disc disease. Decisions concerning coverage for those procedures are made at a local level.

Current Request

On August 16, 2005, CMS accepted a request from Richard A. Deyo, M.D., for an NCD on the CHARITE-Lumbar Artificial Disc Replacement for non-coverage. Dr. Deyo had a concern that, “the Charite disc is a new technology whose real place in spine therapy remains to be determined,” and that important scientific evidence generalizing safety and effectiveness is not available.

CMS is evaluating lumbar artificial disc replacement with a particular focus on the Charite lumbar artificial disc in this analysis, since this was the only disc implant that had FDA approval at the time this national coverage analysis (NCA) was completed. However, we anticipate that when another lumbar spinal disc implant receives approval from the FDA that CMS will, by external request or internal direction, open this NCD for reconsideration with a thorough review of the evidence.

Benefit Category

Medicare is a defined benefit program. An item or service must fall within a benefit category as a prerequisite to Medicare coverage. §1812 (Scope of Part A); §1832 (Scope of Part B); §1861(s) (Definitions of Medical and Other Health Services). LADR with the Charite lumbar artificial disc would be eligible for coverage under Part B, as physicians services, under §1861(s)(1) and (2)(A) and under Part A, inpatient hospital services, under §1861(b).

IV. Timeline of Recent Activities

August 5, 2005 Dr. Richard A. Deyo submitted a letter formally requesting an NCD on the CHARITE-Lumbar Artificial Disc Replacement for non-coverage.

August 16, 2005 CMS opened the NCD process pursuant to Dr. Deyo's request. A tracking sheet was posted on the web site and the initial 30 day public comment period commenced.

Meeting with DePuy Spine.

September 12,
2005

September 16, 2005 The initial 30 day public comment period ended.

March 6, 2006 Meeting with J&J, DePuy Spine.

V. Food and Drug Administration (FDA) Status

The FDA approved the CHARITE™ Artificial Disc in October of 2004 (<http://www.fda.gov/cdrh/pdf4/p040006.html>) .

The CHARITE™ Artificial Disc is indicated for spinal arthroplasty in patients who are skeletally mature, have degenerative disc disease at one level in the lumbar spine from L4 to S1, have no more than 3 mm of spondylolisthesis at the involved level, and have had no relief from pain after at least six months of non-surgical treatment. It is an artificial intervertebral disc made from metal and plastic that is used during a spinal arthroplasty to replace a diseased or damaged intervertebral disc and also treat pain associated with degenerative disc disease.

VI. General Methodological Principles

When making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service falling within a benefit category is reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member. The critical appraisal of the evidence enables us to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve health outcomes for patients. An improved health outcome is one of several considerations in determining whether an item or service is reasonable and necessary.

A detailed account of the methodological principles of study design that the agency utilizes to assess the relevant literature on a therapeutic or diagnostic item or service for specific conditions can be found in Appendix A. In general, features of clinical studies that improve quality and decrease bias include the selection of a clinically relevant cohort, the consistent use of a single good reference standard, and the blinding of readers of the index test, and reference test results.

Public comment sometimes cites the published clinical evidence and gives CMS useful information. Public comments that give information on unpublished evidence such as the results of individual practitioners or patients are less rigorous and therefore less useful for making a coverage determination. CMS uses the initial public comments to inform its proposed decision. CMS responds in detail to the public comments on a proposed decision when issuing the final decision memorandum.

VII. Evidence

A. Introduction

A summary of the evidence used to arrive at the determination is provided. This summary represents the evidence relating to the treatment of pain from degenerative disc disease with LADR with the Charite lumbar artificial disc and includes a clinical trial, case series reports, and technical reviews. The evidence CMS examines has as its focus health outcomes, or, the benefits and harms of a particular treatment. Outcomes that are usually heavily weighted by CMS - morbidity and mortality - are difficult to examine in the context of treatment for chronic low back pain which is a symptom, not a disease. In chronic low back pain, sustained improvement in pain perception and a reduction in the pain-related functional restriction are generally the focus of study outcomes. Measuring a reliable improvement in chronic pain is problematic as pain is subjective and is particularly responsive to the placebo effect; therefore, clinical trials with appropriate controls utilizing independently assessed validated instruments are most heavily weighted. The measurement of treatment effect for low back pain has shifted from physician-based assessment (with outcomes of excellent, good, fair, and poor) to a patient-based self-report of pain and disability (Hagg, Fritzell et al. 2003).

Treatment effect in chronic low back pain is measured with patient-based, multi-item instruments. Two instruments validated for measurement of back pain are commonly used in the assessment of low back pain from degenerative disc disease (Hagg, Fritzell et al. 2003). The Oswestry Disability Index (ODI) is a condition-specific outcome measure used in the management of spinal disorders. The measure is an indication of the extent to which a person's functional level is restricted by pain. The other commonly used measure in chronic back pain treatment effect is the visual analogue scale (VAS), which is a method to assess pain intensity. With the use of these instruments for measurement, a consideration must be given to the clinical meaning of a change in the score (or, for a change in instrument score to be clinically meaningful the patient should experience a change in how he feels or functions). Other considerations include the error of measurement of the instrument used and the clinical importance of a statistically significant score change. In a 2003 study by Hagg of 289 patients treated surgically or non-surgically in a randomized controlled trial, the standard error of measurement of the ODI was 4 units, with a 95% tolerance interval of 10, and the minimum difference that appeared clinically important was 10 units (Hagg, Fritzell et al. 2003). The minimal clinically important difference of VAS back pain was 18 – 19 units with a 95% tolerance interval of 15. It was interesting to note that in this study, improvement after treatment tended to occur to a greater extent in sleep disturbance, ability to do usual things and psychological irritability, but to a lesser extent in the ability to sit, stand and lift.

Some investigators have used the Stauffer Coventry classification, or some modification thereof to measure results. The criteria for clinical results for the Stauffer and Coventry classification are provided in Table 1 (Sott, Harrison 2000).

Table 1

Pain relief (%)	Return to work	Physical restriction	Use of analgesics
Good 76 – 100	Yes	No or slight	No

Pain relief (%)	Return to work	Physical restriction	Use of analgesics
Fair 26 – 75	Yes, with limitations	Yes, limited activities	Frequent (mild)
Poor < 25	No, disabled	Yes, greatly limited	Regular (strong)

Additionally, other quality of life measures are sometimes used. The SF-36 Health Survey, a 36 question form that measures general health status, can be used. Of the 8 health profiles that are included in this survey, only one or two components may be reported, such as the physical functioning composite score or the mental health composite score.

Physiologic segmental mobility is viewed as an important design feature of the artificial disc. Some studies have reported range of motion as an outcome. With fusion, there is reduced motion at the fused segment. The FDA has defined fusion as < 5 degrees of angular motion (FDA guidance document 2000). The theoretical mobility provided by the artificial disc has yet to directly correlate to a proven benefit in how the patient feels or functions, making the clinical significance of post treatment range of motion unclear. Therefore, CMS does not consider post treatment range of motion an important clinical outcome of interest in this memorandum.

Well designed clinical trials can provide the strongest evidence for treatment effect. Clinical trials can be designed to show superiority, a priori, where the superior clinical performance of the investigational agent as compared to the control agent is anticipated. When the investigational agent is believed to have comparable efficacy to the control, but has other advantages, for example fewer adverse events or less cost, a noninferiority trial is an option. In a noninferiority trial, the aim is to demonstrate that the investigational agent is not worse than the control by a certain pre-specified margin, referred to as the delta. In the statistical approach for noninferiority analysis, the delta is compared with the one-sided 95% confidence interval for the difference between the success rate point estimates of the investigational agent and control. If the lower bound of this one-sided confidence interval is less than the delta, then the statistical definition of noninferiority is met.

B. Discussion of evidence

1. Question:

The development of an assessment in support of Medicare coverage decisions is based on the same general question for almost all requests: "Is the evidence sufficient to conclude that the application of the item or service under study will improve health outcomes for Medicare patients?" For this NCD, the question of interest is:

Is the evidence sufficient to conclude that LADR with the Charite lumbar artificial disc will have health benefits for low back pain due to degenerative disc disease in the Medicare population?

2. External technology assessment

CMS did not commission an external technology assessment (TA); however, an external TA was identified on the topic of Artificial Vertebral Disc Replacement.

In April of 2005, the Blue Cross Blue Shield Technology Evaluation Center (TEC) published a TA titled, Artificial Vertebral Disc Replacement. Artificial Vertebral Disc Replacement met only one of five of the TEC criteria. The TEC determined "...the use of artificial vertebral discs for degenerative disc disease does not meet the TEC criteria." The following criteria were not met: 1) The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes; 2) The technology must improve the net health outcome; 3) The technology must be as beneficial as any established alternatives; and, 4) The improvement must be attainable outside the investigational setting.

In summary the TEC report stated the following.

"...the evidence supporting the effectiveness of the Charite artificial disc is limited. Case series provides little evidence of efficacy, particularly in the case of back pain due to degenerative disc disease, where outcomes can be influenced by patient selection, placebo effects, or natural history."

It was further stated:

“The only randomized, controlled trial has several methodologic issues that make it difficult to interpret results.” “...A noninferiority trial design implies that there is a trade-off between efficacy outcomes and some other advantage of a new technology, for example, morbidity or invasiveness, such that a less-stringent threshold for efficacy is acceptable. However, at this time, no such advantage has been demonstrated for the Charite artificial disc.”

Another concern identified was “...that the lack of a prespecified analysis plan, unexplained closure of the database before all patients reached completion, and lack of intent-to-treat analysis may cast some doubt on the analysis.”

The TEC conclusion was, “Given the broader clinical context, and the concerns with the sole randomized, controlled trial, the evidence is not sufficient to conclude that the use of artificial vertebral disc improves health outcomes.” They further concluded that “The evidence is insufficient to determine whether the use of artificial vertebral discs improves net health outcome or whether they are as beneficial as any established alternative.”; and expressed concern “Whether the use of artificial vertebral discs improves health outcomes has not been established in the investigational settings.”

3. Internal technology assessment

CMS performed an extensive literature search utilizing PubMed for new randomized controlled trials (RCTs) and systemic reviews evaluating the use of lumbar artificial disc replacements for the treatment of degenerative disc disease. The literature search was limited to the English language and specific to the human population, but included studies conducted in all countries, including the United States (see evidence tables in Appendix B). Public access information from the FDA website was also used.

Evidence for the Charite lumbar artificial disc came from the FDA PMA Application clinical trial, several case series reports, a systematic review, and adverse events reported to the FDA's Manufacturer and User Facility Device Experience (MAUDE) database.

Evidence Summary

Charite Clinical Trial

In 2005, Blumenthal and McAfee (in two articles) published results of the clinical trial of the Charite III disc compared to the Bagby and Kuslich (BAK) interbody fusion device that led to FDA approval in October, 2004 (Blumenthal, McAfee et al. 2005; McAfee, Cunningham et al. 2005). The investigators' purpose was to compare the safety and effectiveness of LADR, using the Charite disc, with anterior interbody fusion, for the treatment of single-level degenerative disc disease from L4-S1, unresponsive to nonoperative treatment for at least 6 months prior to enrollment. The trial enrolled 375 patients at 14 centers across the United States. The inclusion criteria included: 1) age 18 to 60 years; 2) symptomatic DDD confirmed by discography; 3) single-level DDD at L4-L5 or L5-S1; 4) Oswestry score ≥ 30 ; 5) VAS score ≥ 40 (of 100); 6) failed ≥ 6 mos of appropriate non-operative care; 7) back and/or leg pain with no nerve root compression; 8) able to tolerate anterior approach; 9) able and willing to comply with follow-up schedule; 10) willing to give written informed consent. Exclusion Criteria included the following:

- Previous thoracic or lumbar fusion,
- Current or prior fracture at L4, L5, or S1,
- Symptomatic multilevel degeneration,
- Noncontained herniated nucleus pulposus,
- Spondylosis,
- Spondylolisthesis > 3 mm,
- Scoliosis $> 11^\circ$,
- Mid-sagittal stenosis < 8 mm,
- Positive straight leg raise,
- Spinal tumor,
- Osteoporosis, osteopenia, or metabolic bone disease,
- Infection,
- Facet joint arthrosis,
- Psychosocial disorder,
- Morbid obesity,
- Metal allergy,
- Use of a bone growth stimulator,
- Participation in another study,
- Arachnoiditis,
- Chronic steroid use,
- Autoimmune disorder,
- Pregnancy, or

- Other spinal surgery at affected level (except discectomy, laminotomy/ectomy, without accompanying facetotomy or nucleolysis at the same level to be treated).

Up to five patients at each location (71 total patients) received disc replacement before randomization, which was provided for in the protocol as the opportunity to “insure technical competence with the treatment procedure.” Patients were randomized in a 2:1 ratio (treatment: control), stratified by site in blocks of six. Allocation concealment was done with sequentially numbered envelopes that were opened before surgery, so investigators, staff, and patients were blinded up until this point. Of the randomized patients, 205 patients received the Charite disc replacement and 99 received the BAK cage fusion¹. Control and treatment groups did not differ statistically on gender, age, race, height, previous spinal surgery, and preoperative work status, but did vary on weight (treatment mean 77.5 kg [SD 15.67], control mean 81.7 kg [SD 16.46], $p = 0.0349$), with body mass index being borderline significant (26 [SD 4.23], control 27 [SD 4.76], $p = 0.0557$). The control group received anterior fusion with BAK threaded fusion cages packed with iliac crest autograft. The statistical hypothesis was one of noninferiority, or equivalence, to BAK fusion, where equivalence was defined such that the success rate of Charite was no worse than that of BAK by a pre-specified delta of 15% in the investigator’s protocol. Data were collected at regular intervals up to 24 months post procedure through the trial end and included the ODI, VAS, SF-36 scores, and radiographic information. In a publicly accessible FDA document, concomitant disease of the participants is mentioned as is their activity level before back injury and pre-operatively, as listed in Table 2 (FDA In-depth Statistical Review for Expedited PMA 2004). Though the distribution of normal activity level before back injury did not differ at baseline between treatment and control, there were significantly more active patients in the Charite group than in the BAK group just before surgery ($p = 0.02$). Significance testing comparing treatment and control concomitant disease at baseline was not reported. The reviewer noted, “There were some covariate imbalances between the two groups (e.g., age, BMI and pre-operative activity level, etc...) indicating that patients in the BAK group could have been worse to start.”

Table 2

Baseline evaluation	Charite n= 182	BAK n= 85	p-value *
<u>Normal Activity level</u> <u>Before back injury</u>			
Active	167 (92%)	73(86%)	0.23
Moderate	13 (7%)	10 (12%)	
Light	1 (1%)	2 (2%)	
Minimal	1 (1%)	0	
<u>Pre-operative Activity</u>			
Active	9 (5%)	0	0.02
Moderate	25 (14%)	5 (6%)	
Light	48 (26%)	23 (27%)	
Minimal	100 (55%)	57 (67%)	
<u>Concomitant Disease (> 3%)</u>			
Hypertension	15 (8%)	12 (14%)	
Asthma	11 (6%)	6 (7%)	
Hepatitis	5 (3%)	7 (7%)	
Osteoarthritis	7 (4%)	3 (4%)	
Anemia	6 (3%)	4 (5%)	
Peptic Ulcer	6 (3%)	3 (4%)	
Cancer	2 (1%)	3 (4%)	
Other	80 (44%)	33 (39%)	
* Fisher's exact test for categorical variables and t-test for continuous variables			

Clinical success was judged in terms of a composite outcome that required these four criteria to be met:

1. Improvement in the ODI of at least 25% at 24 months as compared to baseline,
2. No device failures requiring revision, reoperation or removal,
3. Absence of major complications, defined as major blood vessel injury, neurologic damage, or nerve root injury, and
4. Maintenance or improvement in neurological status at 24 months, with no new permanent neurological deficits compared to baseline.

Patient accountability from the Blumenthal article and the FDA summary of safety and effectiveness data is reported in Table 3 (FDA Summary of Safety and Effectiveness Data 2004).

Table 3

12 months			24 months	
Patients	Charite	Control	Charite	Control
Randomized	205	99	205	99
Deaths	1	0	1	0
Failures	7	4	12	8
Withdrawn	5	9	16	17
Expected	192	86	176	74
Missed	8	5	15	8
Actual	184	81	161	66
<p>Note: failures include device removals, revisions, and supplemental fixations. Expected = randomized patients – deaths – failures – withdrawn. Missed are those patients who were “out of the window” of the protocol.</p>				

No statistical analysis plan was in the original protocol documents (FDA in-depth statistical review for expedited PMA 2004). The 71 patients that were regarded as training patients were not included in this analysis. Using the composite measure for clinical success, the artificial disc had an overall success rate reported as 57% and the BAK cage had a success rate of 46%, with the noninferior p value listed as 0.0001. Numerators and denominators were not given for these numbers, but this appears to be a strict intention to treat analysis. The FDA requested that the data also be analyzed and reported using: 1) an improvement in the ODI \geq 15 points at 24 months compared to the score at baseline; and 2) a noninferiority margin of 10%. The FDA concluded that “The two-sided confidence interval indicates that the overall success rate for the Charite Artificial Disc is not worse than the control rate by more than 10%, regardless of which set of study success criteria is used” (FDA Summary of Safety and Effectiveness Data 2004). Though, for the previous analysis the table says “Comparison of Success Rates for Efficacy at 24 months”, the number of subjects (completers) was 184 for Charite and 81 for the control, which corresponds to the 12 month completers. Blumenthal, et al., stated, “Sensitivity analysis were performed to evaluate the potential impact of incomplete subjects (e.g., lost to follow-up).” There was no difference between the two groups as far as operative time (111 minutes in the investigational group and 115.3 minutes in the control group, $p = 0.562$), blood loss (207 cc in the investigational group and 224 cc in the control group, $p = 0.6012$), or level of implantation. Mean duration of hospitalization did differ; however, discharge criteria were not standardized (investigational 3.7 days [SD 1.8], control 4.2 days [SD 1.99]; $p = 0.004$).

In both investigational and control arms the VAS and ODI scores improved at all follow-up times compared to baseline ($p < 0.001$ for all times). Changes were more rapid in the Charite group, and the difference between the two groups was statistically significant until the 24 month follow-up for both groups. No conclusions can be made in respect to time to improvement as the study was designed to demonstrate noninferiority at the 24 month time frame only. At 24 months, the mean decrease in ODI in the investigational group (from 51 to 26) as compared to the control group (52 to 33) did not differ significantly ($p = 0.267$); similarly, VAS decrease in the investigational group (from 7.2 to 3.1) as compared to the control group (7.2 to 3.7) did not differ significantly ($p = 0.1074$). For the component SF-36 scores, 99 (73%) Charite patients and 41 (66%) BAK patients had a 15% or greater improvement in the Physical Composite Score (PCS) at 24 months, and 68 (50%) Charite patients and 34 (55%) BAK patients had a 15% improvement for the Mental Composite Score (MCS). These were not statistically different with a $p = 0.345$ and 0.4959 , respectively (FDA PMA Memorandum Clinical Review 2004).

The adverse events from the FDA clinical review are listed in Table 4 (FDA PMA Memorandum Clinical Review 2004). Device failure appeared to be incorporated into the composite clinical success score, but the category in the composite score for major complications was narrower than the adverse events listed in Table 4 (Blue Cross Blue Shield 2005). Device-related adverse events do not appear to be incorporated into the composite score (Blue Cross Blue Shield 2005). From Table 4, Charite had a higher percentage of patients with severe or life threatening events and device related adverse events; p values were not given.

Table 4*

Adverse Events	Charite disc	BAK cage
Patients with severe or life threatening events	15% (30/205)	9% (9/99)
Device-related adverse events	7.3% (15/205)	4% (4/99)
Device failures	4.9% (10/205)	8.1% (8/99)

Adverse Events	Charite disc	BAK cage
*FDA PMA memorandum Clinical Review 2004		

At 24 months, many patients who met the criteria of clinical success were using narcotics to control pain. In the patients deemed a success at 24 months, the rate of narcotic usage in the investigational group was 64% (73 of 114) and 80.4% (37 of 46) in the control group. Narcotic usage was not defined and was not given for patients who did not meet the success criteria at the 24 month follow-up.

Follow-up radiographs were obtained, scanned, digitalized, and analyzed by a software program (McAfee, Cunningham et al. 2005). At 24 months, range of motion (ROM) as measured on lateral flexion/extension films was 113.6% of preoperative measure (a 13.6% increase from baseline) in the investigational group and decreased in the control group, as would be expected for fusion. No statistically significant association was found between ROM and success/failure at 24 months for those who had data available to the FDA reviewer (FDA Summary of Safety and Effectiveness Data 2004). Fusion rate was judged to be 91.9%. At 24 months, 82.9% of both the Charite artificial disc training and randomized subjects were graded as having ideal placement, 10.7% as suboptimal placement, and 6.2% as poor placement. The authors stated that ODI scores at 24 months correlated with the degree of technical accuracy ($p < 0.05$), as did VAS scores ($p = 0.016$). The disc was more effective in restoring height of the collapsed disc space as compared to fusion ($p < 0.05$), and had less subsidence ($p < 0.05$). The authors stated that, "Long-term radiographic follow-up is necessary to determine if TDR [total disc replacement] can prevent adjacent segment breakdown."

Neurologic status included the following information: reflexes at the knee and ankle; motor function; sensitivity to light touch; strength of lower extremities; and straight leg raise. Neurologic status was reported as equivalent between the investigational group and control group at 6, 12 and 24 months (Geisler, Blumenthal et al. 2004). There was no significant difference between the proportion of patients with neurological adverse events comparing investigational group (16.6%) to control group (17.2%).

Limited information was provided in the FDA's summary of safety and effectiveness data for the initial training group of 71 patients (FDA Summary of Safety and Effectiveness Data 2004). There were higher early (within the first 2 days of surgery) adverse events in this group (33 patients, or 46.5%) than in the randomized group (58 patients, 28.3%). The rates at all other time periods were similar between the two groups. There were more device-related adverse events in the training cases (8 events, 11.3%) than in the randomized group (16 events, 7.8%).

Case Series

In 1994, Griffith published a retrospective review of 93 patients (3 surgeons' experience) with 139 Charite III implants, referred to as "Model III" (Griffith, Shelokov et al. 1994). A single prosthesis was implanted in 53%, two prostheses in 45%, and 2% received three prostheses. One surgeon provided additional data on 58 earlier design prostheses (Models I and II) that were implanted in 49 patients. The authors analyzed data mainly from those patients implanted with the Model III disc, unless it was noted otherwise. Including all data, average age was 43.0 +/- 7.3 years (range: 25 to 59). For the primary study group, the diagnosis was DDD (65.2%), postnucleotomy syndrome (15.0%), internal disc derangement (10.9%), failed fusion (3.3%), instability (1.2%), and herniated nucleus pulposus (1.2%). Forty-one percent had prior back surgery. Average follow-up was 11.9 +/- 8.3 (n= 90) months (range of 1 to 37 months). Three percent of patients who were implanted with Model III were not followed up or were lost to follow up. If patients had more than one follow-up, the last follow-up visit was used for the analyses. Pain experience was measured using a 10 point analog pain scale for right and left leg pain and back pain. There was improvement in the intensity of pain in all three areas ($p < 0.001$), with a total of 71 patients (one surgeon did not report analog pain scores) included in this analysis. Patients' pain was also judged by qualitative change in pain intensity for these three areas as increased, decreased, or unchanged. Here, also, most patients had an improvement ($p < 0.01$). Neurologic weakness (by physical exam) was present in 21 patients (23%) before the surgery. At the follow-up, 17 of 21 patients no longer presented with neurologic weakness. Similarly, there was a 50% reduction in the number of patients with a positive straight leg raising (SLR) test on follow-up exam (left leg positive SLR: 69% preoperative to 35.5% on follow-up; right leg positive SLR: 63% preoperative to 38.7% on follow-up). A comparison of the patients' ability to walk preoperatively and at the most recent follow-up (n= 71, data was unavailable from one surgeon) showed that 39% improved their self-reported walking distance, 2% decreased their walking distance, and 58% remained the same. Subjective, clinical estimates of lumbar flexion and extension showed an increase in both flexion and extension ($p < 0.01$), though the authors noted that the data should be viewed with caution since they were obtained retrospectively and "were not independently assessed with a validated technique." No statistical difference in work status at follow-up could be detected, though work status results differed between the individual surgeons (p value not given). The authors stated, "Inappropriate choice of prosthetic size resulting in implant migration/subsidence or dislocation occurred in 6.5% of the patients in which Model III was used; the incidence of these complications as a function of the number of prostheses implanted was 4.3%." The reported complications that were related to the procedure, but not necessarily to the device's function, included: phlebitis/leg thrombosis (2), injured vein (6), wound bleeding/dehiscence (2), superficial wound infection (1), muscle atrophy (1), urinary tract infection (4), incontinence (3), constipation/defecation difficulty (4), nausea (1), skin paresthesia (1), hematoma (11), hypotension by blood loss (1), retroejaculation (1), and sympathetic sign in left leg (1). Complications considered by the authors as equivocal included: allergy (1), instability "feeling" (2), new paresthesia (1), unspecified neurologic (2), abdominal, leg, thigh, or lumbar pain (10). The Charite III had 3 re-operations out of 93 patients (3%). The re-operation rate for I and II due to complications of implantation was reported to be 5 of 49 patients (10%).

In 1996, Cinotti retrospectively analyzed the follow-up of 46 patients who had the Charite III disc implanted (Cinotti, David et al. 1996). Preoperative diagnosis included disc degeneration in 22 patients and failed disc excision surgery in 24 patients. Thirty-six patients had a single level prosthesis and 10 patients had 2 levels implanted. Post- surgery follow-up ranged from 2 to 5 years, with an average of 3.2 years. The patients had been operated on by one surgeon. Surgery contraindications were degenerative changes of the facet joints identified by CT or MRI, disc degeneration adjacent to a fused area and spondylolisthesis. After surgery, 18 patients wore a corset for 3 months, and 28 patients began exercising within 1 week after surgery (no details). Patients were evaluated by one of the authors who did not participate in the disc replacements. Patients' overall satisfaction, as well as pain sensation, need for analgesics, and the ability to resume work or activities of daily living was reported. Guidelines previously reported by the authors (not detailed in this paper) were used to rate clinical results as excellent, good, fair, or poor.

- Clinical outcome was rated as excellent in 11 patients (24%), good in eighteen (39%), fair in fourteen (30%), and poor in three (7%).
- Patients' satisfaction was reported as a great benefit by 14 patients (30%), a great but not complete benefit by 17 patients (37%), a mild improvement by 12 patients (26%), and no improvement or worsening by three (7%).
- Analgesic drugs were taken occasionally by four patients (9%) and continuously by twelve (26%).
- Resumption of work or daily life activities occurred at the same level in 31 cases (67%) and at a lower level in nine (20%). Four patients, (9%), stated they could not work because of severe back pain and 2 were unemployed.
- Clinical results were rated as satisfactory in 69% of patients (25 of 36) who had single level disc replacement and 40% (4 of 10) in those who had 2 level disc replacement.
- Eight of 17 patients who had unsatisfactory results had a subsequent fusion. Seven of the eight underwent posterolateral fusion with the disc implant left in place, with only three of these patients having satisfactory results at follow-up.

The authors stated, “In the present series, the proportion of satisfactory results (63%) was lower compared with the figures reported for arthrodesis (65 – 90%).” They attributed part of this to the learning curve of the surgery, and that the surgical indications and treatment of patients after surgery changed over time with the surgeon’s increasing experience. Two level surgery was discontinued. They further concluded, “The main cause of poor outcomes appear to be an inappropriate selection of patients undergoing disc replacement...”. The authors commented on important issues related to the surgical approach; because “the placement of the prosthesis into the disc space needs a larger exposure of the anterior annulus compared with anterior interbody fusion”, there is a “greater risk of damaging the big vessels and the sympathetic chain”, with mobilization of big vessels carrying a greater risk of complications in elderly patients. The authors reported a complication rate in patients undergoing disc arthroplasty of 19%, with a comparable rate in those who had anterior fusion of 15%. Other complications are also noted. One patient (2%) had an anterior dislocation of the implant, and 9% had subsidence of the prosthesis into the vertebral bodies, attributed to undersizing of the prosthesis. Interestingly, in patients with a malpositioned prosthesis, there often was noted ossification of the intervertebral space. They noted that, “in case of failure of the prosthesis, there is an intrinsic tendency for the motion segment to undergo fusion.” There was an average vertebral motion of 9 degrees in the sagittal plane at the operated level, while the authors suggested that the prosthesis should have provided a range of motion of 12- 14 degrees in flexion-extension. It was unclear what this meant clinically. Longer term follow-up was recommended to monitor for prosthesis failure, wear of the materials, and loosening of the implant.

In 1997, Lemaire reported on a series of 105 Charite patients with average follow-up of 51 months (Lemaire, Skalli et al. 1997). The average age of the patients was 39.2 years (range 24 – 50 years). Fifty patients (48%) had undergone at least one operation. Clinical results were measured using a modified Stauffer-Coventry rating scale which scored low back pain occurrence, radicular pain occurrence, neurologic deficit, medication use, participation in daily living activities, work status preop and post op, and psychiatric status. Results (Table 5) were measured as relative gain (the authors stated, relative gain “ = absolute gain/maximal gain minus preoperative score”, where the assumption is that this is the same formula as in the 2005 Lemaire article, where relative gain is defined as post-operative score - preoperative score/maximum possible score – preoperative score).

Table 5.

Relative Gain	Result	Percentage of Patients
> 70%	Good	79
60% – 70%	Fair or Satisfactory	5.8
< 60%	Poor	15.2

The authors attributed bad results to incorrect indications (osteoporosis, posterior osteoarthritis, overlying thoracolumbar kyphosis), secondary progression of a posterior facet joint syndrome, or non-return to work. Eighty-seven percent of patients returned to work. Sixty percent had the same work activity, 27 % had reduced activity, and 13% did not return to work. Ten percent returned to “intense sports activities.” Complications occurred in 11 patients (10%): 5 vascular problems (2 phlebitis, 2 pulmonary embolism, 1 acute leg ischemia); 2 temporary neurologic deficits (1 total regressive sexual disorder at 1 year, 1 paralysis at L5 with recovery after revision and fusion); and 4 cases of bone related complications (1 L5 endplate fracture requiring revision with arthrodesis, 1 L5 endplate subsidence of osteoporotic origin, 2 periprosthetic ossifications). Only 3 of these complications were attributed to technique. The results were also analyzed anatomically, biomechanically, and kinematically from X-rays. The average L4-L5 mobility was 9.7 degrees flexion, 3 degrees extension, and 4 degrees lateral bending. At L5-S1 level, these values were 6 degrees, 3 degrees, and 3 degrees, respectively. Interestingly, there was a correlation between posterior joint pain and anterior positioning greater than or equal to 4mm. These authors concluded, “In fact, the proper indication for surgery is crucial for good results.”

A 1999 study by Zeegers in the Netherlands reported on 2 year results for 50 prospectively studied patients that he had operated on (Zeegers, Bohnen et al. 1999). There was a 13% rate of permanent side-effects and/or complications, with 4% related to poor implantation technique. Seventy-five prosthesis were placed in 50 patients: 29 patients had one level insertion, 18 had two level insertion, and 3 patients had three prostheses inserted. Four patients were lost to follow-up (unclear how these patients were regarded). The mean age at the operation was 43 years (24 – 59 years). Mean duration of low back complaints was 10 years (range 1 – 35 years). Fifty-four percent had undergone previous surgery. Patients under 45 years of age were associated with a statistically significant better outcome (< 0.05). Seventy percent of the patients had a positive clinical result defined as a good or fair result from the Stauffer and Coventry criteria. Sixty-five percent (30/46) showed improvement of low back pain. Eighty-one percent returned to some work and 43% returned to their original work. Only fifteen out of 34 patients were able to decrease their analgesic intake. Twelve patients (24%) out of the fifty initial patients needed re-operation which involved 24 procedures (re-operation at the segment with a prosthesis: 6; re-operation at other levels: 11; re-operation related to complications: 7). Side-effects or complications at, or after, the first implantation operation were reported 52 times by 30 patients, with permanent sequela and complications seen in 13%. These included dysaesthesia of legs (3 permanent), painful/numb scar or hematoma (17 temporary), abdominal problems (3 temporary), new or progression of old pain (5 temporary), sympathectomy effect (4 permanent), aortic lesion at removal of prosthesis (temporary), general complication of urinary tract infection, impotence, deep venous thrombosis (5 temporary), malposition of prosthesis (one temporary and one permanent). Fourteen percent of all levels with a prosthesis showed a decrease in height 2 years after surgery. There was no significant migration ($> 2\text{mm}$). The range of motion of the prosthesis between flexion and extension averaged 9 degrees 2 years postoperatively, which equaled the preoperative ROM. The authors provided these comments, “A critical review of our good and poor clinical results makes clear how difficult it is to find the real origin of low back pain,” and, “Several indications and contra-indications for ADR [artificial disc replacement] have been previously reported, but are not unanimously accepted.”

A small case series study by Sott and Harrison in the UK attempted to study patients over 45 years of age (Sott, Harrison 2000). They mentioned that an upper age limit of 45 years was proposed by the manufacturers and several authors, “as increasing age may lead to weakening of the bone structure supporting the prosthesis.” Fifteen prostheses were implanted into 14 patients aged 31 to 61 years (mean age 48 years). Nine prostheses were implanted at level L4/L5, four at L3/L4 and two at L5/S1. None of the patients were able to carry out their usual professional, domestic or leisure activities without pain preoperatively. The patients were followed for an average of 48 months (18 to 68 months). Four patients were followed up by phone. Criteria for clinical results were according to the Stauffer and Coventry classification. The patients were divided into two groups, those less than 45 years (7) and those over 45 years (7). Patient outcomes related to age were identical for both groups: 5/7 good, 1/7 fair, and 1/7 poor. There was one case of implant migration in a woman with normal bone density preoperatively. Another patient required fusion for symptoms related to a non-operated level.

A 2002 abstract by David reported on 147 patients implanted with the Charite prosthesis with a minimum of five years follow up (David 2002). Patients had 163 prosthesis implanted (16 at two levels) in L4-L5 and/or L5-S1 for chronic low back pain alone (59), or with sciatica (88). Seventy two patients had been operated on before. The results were stated as 79% of 142 patients had excellent or good results using the Stauffer Coventry classification. One patient had removal with fusion for severe sciatica, two had secondary bone migration with fusion and ten patients had fusion (prosthesis left in place) for malpositioning and facet pain. Eleven patients had partial or total ossifications around the prosthesis. The author concluded, "Good positioning of the implants is very important for long term mobility to avoid facets deterioration."

In 2003 a systematic review of case series studies was reported by de Kleuver (de Kleuver, Oner et al. 2003). This review included 6 of the 7 short term case series [246 patients] mentioned in this document (all except Caspi 2003), two foreign language articles [74 patients], and 6 Acroflex disc patients. The authors reported that patients classified as having "good" or "excellent" results varied in the studies from 50% to 81%. Various complications were observed in 3-50% of patients, including vascular injury, implant migration/subsidence, or dislocation. A meta-analysis could not be performed due to the lack of comparative studies. The authors concluded that there was insufficient data to assess the performance of total disc replacement.

In 2003, Caspi reported the outcome of 20 patients implanted with Charite III after a 48 month follow-up (Caspi, Levindopf et al. 2003). Preoperative diagnosis included degenerative diskopathy (DDD) in 17 patients and failed posterior conventional discectomy in 3. Seventeen patients had one level implantation and 3 patients had two level implantation. Age range was 24 to 50 years. Three of 20 patients had undergone previous surgery by a posterior approach. The authors stated the results of these 20 patients as:

“The overall clinical results were rated as follows: fair = 3, good = 4, excellent = 11, and poor = 4 (one patient underwent secondary fusion and one is waiting for surgery). With regard to the patients’ recovery in terms of occupation: four are completely disabled, one patient resumed physical labor, and the others returned to light and sedentary work.”

There were two cases of migration of the prosthesis, one intraoperative laceration of the ureter and thrombosis of the iliac artery occurred. Two patients had ossification of the intervertebral anterior ligament. Average range of motion was given as 3 – 9 degrees. Though it was stated that radiologic results were analyzed from X-rays and that the clinical outcome was assessed by comparing presurgical with follow-up data using the Oswestry questionnaire and the visual pain analogue scale, no details were given as to how the statement of clinical results was arrived at. One of the conclusions of this article was, “Contraindications for surgery appear to be the principal cause for failure rather than the prosthesis itself.”

In 2005, Lemaire reported on a 10 year minimum follow-up of 100 of 107 original patients (Lemaire, Carrier et al. 2005). They were followed for a minimum of 10 years (range 10 – 13.4 years). Seven patients were unavailable for long-term follow-up. Fifty-four patients were implanted with one-level, 45 were implanted with two-level, and one with three-level prostheses. Age at time of surgery was 39.6 years (range = 23.9 – 50.8 years). While the mean age and range were very close to the 1997 Lemaire study, the mix of male and female differed (1997 [68 M, 37 F], 2005 [41 M, 59 F]). Patient indications included DDD with low back pain of discogenic origin at one or two levels (one had three-level surgery) and failure of nonoperative treatment (PT, medication, exercise). Contraindications were obesity, prior fusion, instability, deformity, radicular pain symptomatology, and facet arthrosis. Radiographs, MRI (not in early cases), and provocative discography were done preoperatively. Clinical outcomes were determined using a modified Stauffer Coventry scoring system, as in the previous Lemaire, et al., 1997 study. The results were reported as relative gain in the modified Stauffer Coventry Scale, which was defined as post-operative score - preoperative score/maximum possible score – preoperative score, expressed as a percentage. A relative gain of at least 70% was defined as excellent (no pain, no medication, resumption of activity in the same job after 3 months), 60 – 69% as good (intermittent and infrequent lumbar pain not requiring major or prolonged medication, resumption of activity in the same job after > 3 months or in a less strenuous job after < 3 months) and less than 60% as poor. By this definition, 62% of patients had excellent results, 28% good results and 10% poor results. For the 95 patients who had not retired, 87 returned to work. There was no statistical difference in outcomes between patients with one level prosthesis versus two level ($p < 0.05$). Radiologic analysis was performed using anatomic, biomechanical and kinematic criteria. By radiologic evaluation, no subluxation of the prosthesis or the core was noted. Minor subsidence was noted in two patients, both due to trauma. Osteolysis was not identified. Sixteen of the 59 women were menopausal at the time of follow-up, with no radiologic signs of osteoporosis. Pelvic anteversion (increase of the sacral slope, reduction of the pelvic version) was noted as being “routinely obtained or preserved”. Mean range of motion was reported as 10.3 degrees flexion/extension and 5.4 degrees lateral bending. Five patients (5%) had a secondary arthrodesis, with two of the five reported as having good outcomes. Four had symptomatic articular arthritis. Two patients (3%) had periprosthetic ossification (bone formation adjacent to the prosthesis) affecting prosthesis mobility, which the authors compared to 7.7% in David’s series (David 2002) and 1.7% in Marnay’s series (Marnay 2002). They reported no correlation between ossification and outcomes. If the ossifications were lateral, the segment fused, if anterior, the segment was mobile. Lemaire, et al., stated, “Because these ossifications appear after the fifth year postoperatively, the risk of adjacent functional overload can appear as late as after the 15th year postoperatively.” Two patients had adjacent level degeneration. The level of identified complications (9%) was reported as being equal to David’s experience, and lower than Marnay’s Prodisc series (26%).

In a 2005 article Putzier reported on clinical and radiographic results of 17 years from Charite – University Medicine in Berlin (Putzier, Funk et al. 2005). It was a study of 71 consecutive patients treated surgically with either Charite I, II, or III between 1984 and 1989 with 84 Charite disc implants. The follow-up averaged 17.3 years (14.5 – 19.2 years), and ultimately 53 patients (74.6%) were available for examination. There were 20 males and 33 females with an average age of 44 years (30 – 59 years). Treated levels included L3/4, L4/5, L5/S1, and L4-S1. Fifteen patients had type I, 22 patients had type II, and 16 had type III. Surgical indication was one or 2 segment DDD of the spine. Eight patients had had previous disc surgery, and 3 had spondylolisthesis grade I. All operations were performed by experienced senior spine surgeons and included the designers of the prosthesis. Clinical examination included the ODI and the VAS. The patient's perception of the overall outcome was graded at follow-up as excellent, good, fair, or poor according to Odom's criteria (a functional outcome scoring system). Radiological parameters included plain X-rays with flexion/extension views. A segmental mobility of 3 degrees or less was graded as fused, with more the 3 degrees graded as mobile. Adjacent segments were evaluated to determine any progression of degeneration in comparison to preexisting X-rays. Clinical results are as follows:

Oswestry disability index Type I: mean 40.76 SD 19.64

Type II: mean 45.67 SD 22.45

Type III: mean 37.18 SD 22.12

Visual analog scale Type I: mean 40.76 SD 19.64

Type II: mean 45.67 SD 22.45

Type III: mean 37.18 SD 22.12

Outcome criteria according to Odom, in relation to the type of prosthesis

	excellent	good	fair	Poor	GPA*
Type I	27% (4)	27% (4)	33% (5)	13% (2)	2.33
	23% (5)	32% (7)	13% (3)	32% (7)	2.55

	excellent	good	fair	Poor	GPA*
Type II					
Type III	31% (5)	27% (4)	33% (3)	13% (4)	2.38
*grade point average, with excellent grade 1, good grade 2, fair grade 3, poor grade 4.					

Upon statistical examination, there was no significant difference between the three types of Charite for all three clinical parameters. Radiological results revealed that of the 53 patients available for follow-up, 12 (23%) had surgical fusion. Of the remaining 41 patients, only 9 patients (17%) did not have heterotopic ossification of ankylosis.

	Classification of heterotopic ossification/fusion *				
	0	I	II	III	IV
Charite disc type					

	Classification of heterotopic ossification/fusion *				
I	0 (0%)	0 (0%)	0 (0%)	0 (0%)	12 (80%)
II	0 (0%)	0 (0%)	2 (9%)	1 (5%)	11 (50%)
III	1 (6%)	1 (6%)	2 (13%)	2 (13%)	9 (56%)
Total	1 (6%)	1 (2%)	4 (8%)	3 (6%)	32 (60%)
*Classification of heterotopic ossification/fusion according to McAfee					

Adjacent segments were evaluated for degenerative changes, with 9 cases having significant degenerative changes. The authors commented, “Adjacent segment alterations were found only in cases where spontaneous ankylosis of the treated segments, spondylodesis, or fusion after implant failure occurred.” The different types of discs were not found to influence this (no statistical significance), nor did it matter if the ankylosis was spontaneous or there was surgical fusion. The authors also correlated clinical parameters with radiological results. Interestingly, the clinical parameters did not correlate with the radiologic evaluation of functional status of the surgically treated segment.

Table 6

Status of segmental Fusion	Patients (n)	Oswestry disability index		Visual analog scale		Odom's Criteria
		Mean	Standard deviation	Mean	Standard Deviation	GPA
No fusion	9	52.09	14.42	6.08	1.35	2.67
Spontaneous Ankylosis	32	37.84*	10.41	4.45*	1.14	2.34
	12	44.28	17.04	4.93	1.56	2.50

Status of segmental Fusion	Patients (n)	Oswestry disability index		Visual analog scale		Odom's Criteria
		Mean	Standard deviation	Mean	Standard Deviation	GPA
Fusion after implant Failure						
*p < 0.5 in comparison to “no fusion”						

Adverse Events

van Ooij, in 2003, reported on a series of 27 patients who presented to a tertiary care center with an unsatisfactory result or complication after Charite disc replacement (van Ooij, Oner et al. 2003). These patients, (except for one), belonged to a series of approximately 500 patients operated on in a single institution. The mean age was 40 years (range 30 – 67 years) at the time of operation. Presentation was at a mean of 53 months (range 11 – 127 months) following total disc replacement. Twenty-two patients had the prosthesis implanted at a single level, four patients received two disc implants, and one patient received three implants. Early complications included one patient who had a dislocated disc implant anteriorly within one week postoperatively, then had a subsequent fusion cage placed, and continued to have disabling back pain 2 years out. The 26 late complication patients were described as, “The clinical picture of most of the patients was of a very disabling nature.” Patients often had a combination of pathologies. Degenerative disc disease at another level (either present before the operation or developed afterwards) was seen in 12 patients. Facet joint arthrosis (at the operated level or at a neighboring level) was seen in 11 patients. Subsidence was present in 18 patients. Two patients had migration of the prosthesis which resulted in compression on the great vessels in one patient. Breakdown of the polyethylene was seen in one patient.

Adverse events reported in the Manufacturer and User Facility Device Experience (MAUDE) database

The FDA provided an analysis of adverse events reported in MAUDE database at the request of CMS (FDA memorandum to CMS 2005). The analysis includes Medical Device Reports (MDRs) that were entered into the database between August 11, 2003 (date first report was received) and November 16, 2005. A total of 101 MDRs were analyzed for 96 patients, 1 MDR for the Prodisc device in addition to the Charite devices. The most frequently reported event was device migration out of the implanted location, with 54 of 96 patients (56%) experiencing this adverse effect. Seventy-six patients (79%) had a second surgery to remove all or part of the implant, to correct problems with the device, or to correct problems produced during the implant surgery. Fifty of the 76 (66%) patients had second surgery due to device migration. The most common second surgery was removal of all or part of the artificial disc followed by spinal fusion of the implanted motion segment. Twelve patients had two prostheses placed despite the device labeling for only one device implantation. Most adverse events that required second surgery occurred in the first 2 months after implantation. Two deaths were reported which were both attributed to pulmonary emboli.

Lumbar artificial discs not yet FDA approved

Peer-reviewed articles on studies of artificial discs not yet FDA approved were referenced during the public comment period for the proposed decision memorandum. However, as this analysis has focused on the Charite disc since it is the only one with FDA approval, we will not review them here.

4. Medicare Coverage Advisory Committee (MCAC) Meeting

The MCAC was not held for this topic.

5. Evidence-based guidelines

No evidence-based guidelines were identified.

6. Public Comments

During the initial comment period, CMS received comments from five national professional societies and 138 public comments. The majority of comments received asked that CMS deny the request for non-coverage of the lumbar artificial disc. Though in this group there was agreement that the request should be denied, opinions varied on patient selection and whether there is adequate evidence for evaluating health outcomes of the lumbar disc arthroplasty in the Medicare population. Those comments and the complete summary can be found on our [website](#).

Comments on the Proposed Decision Memorandum

CMS received a total of 604 comments during the public comment period for the proposed decision memorandum. Of the 604 comments, 129 were posted to the CMS website during the public comment period. CMS received the remaining 475 through the mail or by e-mail and scanned and subsequently posted them to the CMS website. Of the 604 comments, 7 supported the CMS decision for non-coverage, 596 opposed the decision and one comment was unclear on position.

Of the 475 comments not submitted to the CMS web site, 470 were copies of a form letter provided by Texas Back Institute requesting approval for coverage and signed by patients, family members and others. Twenty-five of the 470 form letters included personal comments by the signers. Of the 25 form letters with comments, 13 were from people who stated they had the procedure and 6 were from people planning to have the procedure. The form letter did not reference the content of the proposed decision memorandum other than the statement, "...the conclusion not to cover CHARITE™ Lumbar Artificial Disc Replacement surgery as a "reasonable and necessary" procedure." The form letter stated, "CHARITE™ surgery as an alternative to more commonly performed spinal fusion surgery allows for faster recovery times, more flexibility and increased range of motion without the need for prolonged bracing, while fusion surgery can put additional stress on discs above and below the fusion site that may cause future back problems especially in elderly patients suffering from osteoporosis."

CMS appreciates testimonials from patients submitted as part of its coverage process. However, testimonials are rarely helpful in evaluating the evidence for or against a specific item or service. Moreover, we are skeptical of form letters as they may not always reflect the complete views of the signatory, but instead reflect the opinion of the individual or entity that drafted the form letter in the first place. In this case, it is unclear if any of the respondents were Medicare beneficiaries.

Three additional hard copy comments and 2 e-mails were also received. Of these 5 comments, one was from an advocacy company for the biomedical industry who favored coverage, one was from a physician who attached a document written by a patient seeking coverage for the procedure, one was from a back patient supporting coverage, one was from the spouse of a patient who had the procedure supporting coverage and one was from a patient who had the procedure and is experiencing post operative problems and opposes coverage at this time.

Of the 129 comments posted to the CMS web site, 4 were from 3 national professional societies, 79 were from physicians, 13 were from patients that had the procedure, and 9 were from patients planning on having the procedure. Additional comments were from or on behalf of orthopedic device manufacturers, an employee of a device manufacturer, a trade association, a company representing members of the health care industry in legislative matters, representatives of a health insurance company, local professional societies, patients with back problems, family members of back patients, and the general public. Of these 129 comments, 122 disagreed with the CMS proposed decision, 6 agreed with the proposed decision and one comment was unclear as to position. Of the 79 physicians who posted comments, 75 disagreed with our non-coverage decision and 4 supported our decision. One physician commenter stated, “Depuy Spine, the manufacturer of the Charite artificial disc, is currently orchestrating an aggressive “letter writing” campaign asking surgeons to write CMS and request that coverage be granted. After receiving an e-mail from the company, I felt compelled to write and voice my strong support for non-coverage. I believe that the Depuy strategy is self-serving and is clearly intended to bolster the device’s stagnant sales figures.”

The majority of comments that supported coverage did not provide specific reference to new evidence that was not already part of the NCA. However, 18 commenters made specific reference to evidence.

A. Professional Societies

Four comments from three national professional societies were submitted. Comments were submitted from the American Association of Neurological Surgeons (AANS)(the President), the Congress of Neurological Surgeons (CNS)(the President), the AANS/CNS Section of Spine and Peripheral Nerves(the Chairman), and the North American Spine Society (NASS)(the President). The letters were identical. They suggested that in patients under the age of 65, the evidence is sufficient to conclude that health outcomes will be improved, as well as in properly selected patients in the Medicare age group: “In the population group studied, the data clearly showed that artificial disc replacement was as effective as other current procedures, and therefore should be an available treatment option for Medicare beneficiaries under the age of 65.” They stated, “As with any surgical procedure, careful patient selection is essential and the surgeon, in consultation with the patient, is the best person to decide if his or her patient is a candidate for artificial disc surgery, regardless of the patient’s age,” and, “At present, there is not enough available data on patients over the age of 60 to demonstrate that this procedure is inappropriate for elderly patients.” A recently published small study in the over-60 age group is mentioned (Bertagnoli, Yue et al. 2006a).

CMS welcomes comments from professional societies. One article (on a lumbar artificial disc not currently approved by the FDA) was referenced. Also, contrary to the commenter’s suggestion, CMS has not limited the finding of noncoverage in its NCDs to only those instances in which evidence has clearly demonstrated that the item or service is harmful to the Medicare population. On the contrary, as discussed in Appendix A, under §1862(a)(1)(A) no payment can be made unless an item or service is reasonable and necessary for the diagnosis and treatment of an illness or injury or to improve the functioning of a malformed body member.

CMS does note that the professional societies expressed, “that artificial disc replacement was as effective as other current procedures, and therefore should be an available treatment option for Medicare beneficiaries under the age of 65”. However CMS also notes that the BAK cage treatment has fallen out of favor and may not be considered a current treatment. Given the particulars of the Charite PMA trial, the evidence is insufficient to conclude that LADR with Charite is as effective as other current procedures for the Medicare under age 65 population. It is also noted that the study was limited to those age 60 and under.

B. Expert Opinion

Drs. Blumenthal, Geisler, and McAfee, the primary authors of the FDA IDE trial, provided comments. They did not support the decision by CMS. Two of the 3 provided reference to evidence in their comments which is addressed below. Dr. Blumenthal stated, “While I would agree that this technology has limited use in the over 65 Medicare population, the indications for the 60 and under age group I believe are quite sound” and expressed the sentiment that the decision, “could be used by private carriers as an excuse to deny coverage to this FDA-approved device.” While aware of the potential of its decision to be used in support of decisions by other organizations, CMS bases its coverage decision on the statutory standard, not to influence payment decisions by private carriers.

Dr. Geisler suggested coverage with these restrictions:

- “1) A DEXA scan to verify bone quality and eliminate osteopenic/osteoporosis patients;
- 2) Flexion/Extension radiographic [sic] with the standard radiographic of the lumbar spine (obliques, AP, lateral)
- 3) a CT of the lumbar spine with contiguous fine cuts to assess the facets and pars bony anatomy in detail.”

With this, he stated “there are a decreasing number of patients over forty years of age that would qualify for the Charite artificial disc.” Additionally, he has reanalyzed the FDA IDE data using non-parametric statistics and found that the “results demonstrate superiority of disc arthroplasty over fusion in indicated patients using these two key clinical measures [VAS and ODI] and a major improvement from baseline in both treatment groups.” He cited that some of the case series and reviews have had mixed results (Griffith, Shelokov et al 1994; Cinotti, David et al. 1996; Zeegers, Bohnen et al. 1999; van Ooij, Oner et al. 2003; Putzier, Funk et al. 2005; de Kleuver, Oner et al. 2003; Gamradt and Wang 2005) while other results have been more favorable (Lemaire, Skalli et al. 1997; David 2000; David 2004; Lemaire, Carrier et al. 2005). He noted that, “If this historical review of the prior series had not occurred, then the clinically superior results presented in this paper [FDA IDE study] would not have been possible.”

In response, one of the goals of the NCD process is to clearly define those patient populations for which there is a clear health benefit. Dr. Geisler has proposed interesting criteria that may in fact more accurately identify appropriate patient populations, though these criteria need to be tested through well designed clinical studies. It is noted that he strongly supports the indications for the 60 and under age group.

Dr. McAfee referenced the reanalysis by Dr. Geisler of the PMA data. In this study, the 71 nonrandomized training cases have been included. He stated, “With the “training cases” added the conclusions of the new analysis using the Wilcoxin Rank Sum Test are even more convincing – Significantly Superior Clinical Outcomes at all time intervals through 2 year follow up.” He also referenced the DePuy Spine complication rates from the MDR (see DePuy Spine public comment) and that they are actually lower than the incidence in the FDA randomized portion of the trial. He commented on the Continued Access portion of the FDA study: “In the Charite IDE from the randomized study of 375 to 688 (CA1 + CA2) patients the reoperation rate went from 4.9% up to 8.1% (Compared to 12.1% for Fusion Control).” He also mentioned an analysis that will be presented at IMAST and Scoliosis Research Society Meeting this summer, “Predicted 5-year Survivorship of the CHARITE Artificial Disc vs. Anterior Lumbar Interbody Fusion: A Kaplan-Meier Analysis.”

The following additional information sources that were not discussed in the proposed decision memorandum were provided by Dr. McAfee and Geisler. Gamradt (Gamradt, Wang 2005) was a review of lumbar disc arthroplasty and does not provide additional information, but in reference to the Charite III it is stated, “In most published series, this device is being implanted in patients near 40 years old. In vivo failure rates, long-term pain relief and revision options are therefore critical unanswered questions, necessitating caution when using this device outside of a clinical trial.”

The 2000 abstract by David reported 5 year follow-up of 96 patients. Seventy-five percent of 92 patients (4 lost to follow-up) “were excellent or good among the Stauffer Coventry classification”, with 5 complete ossifications around the prosthesis and 10 patients with secondary fusion judged as bad results. He also commented, “Indications must be rare in balance with fusion which is always possible.”

The 2004 abstract by David reported on 272 (197 patients) Charite disc prostheses with minimum 10-year follow-up. Here, the outcome measures were radiographic results (80% judged as good, no detail). Eighteen patients had instrumented posterior fusion, 10 had complete ossification, 4 had core failure, 4 more appear to have had subsidence (3 of these had posterior fusions), but only 5 had adjacent level problems of which 2 involved disc herniation and 3 involved stenosis with subsequent fusion performed in 2 of these cases. Limited details were provided. He did state, “The best indications are: in balance with fusion, and young and active patients with severe discogenic low back pain.” and, “Functional results are better than, and the incidence of adjacent level problems are much less than, those reported for fusion in the long-term.”

An editorial (McAfee 2005) on the van Ooij article was provided as evidence, in which Dr. McAfee stated, “The key to constructive surgical education examining the adoption of an innovative surgical technology is that one needs to be able to distinguish between complications due to poor surgical technique and complications arising from the inherent shortcoming of the device itself.”

The Hagg 2004 article on predictors of outcomes in fusion surgery for chronic low back pain is referenced, in which the authors concluded, “...that improved selection of successful surgical candidates with CLBP [chronic low back pain] seems to be promoted by attention to severe disc degeneration, evaluation of personality traits and shortening of pre-operative sick leave.”

The Kuslich and Ulstrom 1998 study on Bagby and Kuslich method of lumbar interbody fusion (BAK cages) is described for 947 patients. The 24 month results were reported as, "Fusion occurred in 91% of patients at 24 months after surgery, and pain was eliminated or reduced in 84%. Function was improved in 91%." This is the device that was used in the control group in the PMA Approval trial for the Charite disc. This procedure has largely fallen out of favor.

The panel transcript from the FDA Orthopedic and Rehabilitation Devices Panel of the Medical Devices Advisory committee was referenced. The panel discussion about the Charite disc was wide ranging and some of the concerns expressed were about the ability to assess pain, the study's ability to claim that the device maintains motion, and the lack of evidence that it delays adjacent segment disease. It was highlighted that this topic is very difficult due to the complexity of the disease. Panelists generally felt that there should be a surgical alternative to fusion. The panel voted to approve the device with certain conditions, such as a training requirement and post-approval studies, because the long-term safety and effectiveness of the device was unknown.

Dr. McAfee cited a DePuy supported publication, of which he is the editor (Dr. Geisler and Scott-Young are the co-editors, with all three acknowledging that they are consultants for DePuy and receive financial, material, or grant/research support from DePuy): *Roundtables in Spine Surgery, complications & revision strategies in lumbar spine arthroplasty*. This 213 page monograph has seven sections concerning the complications and revision strategies for lumbar spine arthroplasty and concludes with a roundtable discussion of experts. Dr. Scott-Young authors a section on lumbar disc revision surgery. He noted the difficulty of back pain diagnosis, "The diagnosis of internal disc disruption or facet arthropathy seems to have no widely accepted reference standard, based on the lack of difference in morphologic changes between normal volunteers and those who have the symptoms." He anticipates a rise in the rate of artificial lumbar disc revision procedures based on experience with other surgical procedures: "Revision arthroplasty of the knee and hip currently constitutes 9% to 11% of all arthroplasty procedures in Australia," and that, "If [sic] follows that as the use of spine arthroplasty increases, we will see similar amounts of collateral damage arise." Upon review of the disc arthroplasty literature he stated, "Repeat surgery generally is associated with a less-favorable outcome than the initial procedure."

C. General Public Comments with Evidence

Twelve comments from the general public provided specific reference to evidence. Two device manufacturers, DePuy Spine and an attorney representing Stryker, provided extended commentaries. Italicized headings summarize the comments, followed by an elaboration and the CMS response.

1. These comments by DePuy were in reference to the Charite PMA noninferiority study.

Reporting of outcome results

- The minimum clinically important difference of the ODI and VAS back pain is, “compared to baseline, **not** between treatment arms.”
CMS contends comparing the outcomes to baseline and not between treatment arms was a change in the analysis plan after the completion of the trial. A sound study design includes a predefined analysis plan and post-hoc changes to the analysis plan are inappropriate and limit the usefulness of the data.

- The outcomes in both VAS and ODI far exceeded the minimum clinically important difference at 24 months. CMS again notes that this was not part of the proposed analysis plan. Firstly, VAS and ODI are listed by mean without reporting the variance. It is difficult to know how to interpret the mean without understanding the variability about that point. Secondly, the trial design was that of noninferiority. In this noninferiority design, the Charite lumbar artificial disc is no worse than the BAK cage. Comparing trial end scores to baseline illustrates a methodological issue in this noninferiority trial: “A noninferiority or equivalence trial requires that the reference treatment’s efficacy is established or is in widespread use so that a placebo or untreated control group would be deemed unethical” (Piaggio, Elbourne et al 2006).” Spinal fusion for the indication of discogenic pain is controversial with conflicting results from clinical trials. Advocates of surgery generally cite a single randomized controlled Swedish trial, while critics cite difficulties with that trial and the results of a more recent trial. Deyo, Nachemson, et al., 2004 stated, “Fundamental problems plague the study of spinal fusion, including the lack of definitive methods to confirm a solid fusion, a weak association between solid fusion and pain relief, and the placebo effect of surgery for pain relief.” The BAK cage is not currently in widespread use, and the use of fusion is variable with unclear indications, and thus the noninferiority trial design is inappropriate to demonstrate a health benefit.
- Geisler, et al., has reanalyzed the data from the Charite IDE trial and has discovered the superiority of lumbar arthroplasty as compared to fusion at 2 years (Exhibit B, DePuy Comment). The claim is that the ODI and VAS scores followed a normal distribution at baseline but at the two-year follow-up the distribution of these scores was “somewhat skewed” so a non-parametric test is more appropriate. This analysis includes the 71 subjects that were non-randomized. It was stated, “Further, this secondary analysis provides strong evidence that in fact, subjects receiving the Charite artificial disc experienced improvement in pain and disability level compared to their baseline levels that was both statistically significant and clinically important at all time points through two-year follow up.” Attempting to reanalyze data collected in a trial that has methodological issues is challenging and discouraged by methodologists. Also, including nonrandomized patients with randomized patients reintroduces the selection biases that randomization helps eliminate. Exploratory data analysis is an important part of the data analysis plan. In the initial exploration of data results, the distribution of the data is examined to check for a normal distribution, where the most prominent features are symmetry and “bell shape” of the curve when the data is examined graphically. Skewness is one of the indicators of the shape of the distribution of the data. Considered in isolation, skewness may not be useful in testing for normality as it is vulnerable to outliers. Other formal tests are generally done such as the careful analysis of residuals to confirm the adequacy of the model. Given the premise of the reanalysis, the published article on the PMA trial does not state that this model checking was done in original analysis, though the assumption would be that it was. If it was the addition of the 71 nonrandomized patients that increased the skew which lead directly to the reanalysis, the outliers should be identified and investigated for explanation of these observations that apparently deviate from “usual” results.

- The commenter stated that the Charite disc is safe for its intended use. Depuy analyzed data from MDRs submitted to the FDA to develop rates. Additionally, they conducted a survey of all surgeons who have been trained to implant the device at the DePuy Spine sponsored training courses beginning December 2004. They reported a 13% response rate and stated that “it captured a relatively large proportion; over 30% of the cases performed post FDA-approval.”

Both the data in the FDA MDR database as well as the data referenced here all suffer from lack of a long-term follow-up. As we have discussed previously, LADR is a long-term procedure and safety data is needed after long-term use. In addition, both the FDA MDR database and this survey suffer from significant bias in response rates. It is well-recognized that the FDA MDR database underreports complications. A survey that has a 13% response rate representing 30% of the surgeries also likely underreports complications in that surgeons not responding have a lower volume experience and thus more likely to have complications. The table provided from the survey in the response for adverse events rate is difficult to interpret and does not resemble tabulations in the FDA clinical review tables. Also, second surgery rate (revisions) is not provided (second surgeries were provided in the analysis of adverse events by the FDA for CMS). Thus, this new information does not change our concerns about the long-term safety of this device.

- The commenter took issue with CMS noting that the calculated success rate appeared to be a strict intention to treat analysis. They noted, “The final dataset reported to FDA, and used to prepare the Blumenthal and McAfee manuscripts did not specify an intent to treat (ITT) analysis, however, the principles were adhered to.” CMS contends that, in Blumenthal, numerator and denominator were not given for this number, and in an earlier analysis at 24-month endpoint, randomized subjects who were overdue for their 24-month evaluation and patients who had not reached 24 months’ evaluation were excluded from the analysis which gave a somewhat better success rate (63%), but was a violation of the intent-to-treat principal (Blue Cross Blue Shield 2005). Therefore, CMS clarified that the success rate that was presented in Blumenthal, after calculation check, was a strict ITT, contrary to earlier reported results.

A statistical analysis plan (SAP) was included in the original protocol

- DePuy further elaborated, “The analysis plan was documented in the SAP that was prepared in October 2001. Although the analysis plan was not prepared prior to the start of enrollment it was written before enrollment was completed in April of 2002.”

As the statistician stated in the FDA in-depth statistical review, “There was no statistical analysis plan (SAP) proposed in the original IDE protocol (file versions, Vol.14: Pages 9-180). Only after the Agency’s request during the pre-PMA meeting on Nov. 14, 2003, the sponsor’s consultant statisticians (Kathie Drouin and George DeMuth) from Stattech Services, LLC provided the SAP via email on Nov. 25, 2003. The following analysis plan was on this SAP, dated March 27, 2003, but appeared to be originated on Oct. 18, 2001 and finalized on Oct. 2003. Please note that it is not clear to me when this SAP was proposed and whether the statisticians at Stattech Services, LLC were blinded to the data access before developing the SAP.”

A proper sensitivity analysis was conducted to evaluate the potential impact of incomplete subjects

- DePuy asserted that, “The Proposed Decision Memorandum inaccurately states that the only imputation method was “last value carried forward.” However, a variety of sensitivity analyses were performed in support of the non-inferiority claim.” They further stated, “Again, each of the results, with the exception of the worst-case scenario, supported the non-inferiority claim with $\delta = 0.1$,” and “The PMA approval by the FDA is evidence that the worst-case scenario does not bear merit.”

The Proposed Decision Memorandum did not state that the last value carried forward (LOCF) was the only method of imputation, but that it was a technique that was used. Upon our review of the FDA in-depth statistical review and clinical review it appears that the sponsor did use LOCF.

BAK is a valid comparator

- Several reasons were given as to why BAK was chosen at the time of the trial. Additionally DePuy stated, “Geisler has performed a meta-analysis of clinical results in the literature with more recent interbody devices as well as 360 fusion. This analysis showed that outcomes have not dramatically improved compared to the results with BAK.” CMS contends that fusion results can be quite variable. The meta-analysis by Geisler is presented in two tables in the article where the primary focus is the neurological complications in the Charite IDE trial. Twenty-nine studies of fusion are included (it is not clear if some studies overlap patients). The analysis does not mention if the review was systematic, the inclusion/exclusion criteria of studies, or the degree of heterogeneity in the studies (describes how different the studies are to decide if you can combine the results). The BAK cage is not currently a well-accepted procedure. In addition, as we discuss in the Analysis Section, we have fundamental concerns about the adequacy of the evidence for fusion surgery in general.

In addition, DePuy presented as evidence a letter to the editor by Dr. Buttner-Janz, one of the inventors of the Charite disc. Dr. Buttner-Janz's criticisms of the Putzier article included the following observations.

- “There were 13 patients implanted with the SB I prosthesis, not 15.” Further, “These 13 patients (5 males, 8 females) received 14 prostheses (not 16).” While Dr. Buttner-Janz stated the statistical calculations are flawed, there is no indication of what the revised calculations would be.
- “However, with a lost to follow-up rate of 25.4%, incorrect patient/prosthesis numbers and extremely small sample sizes in each group, the power of their statistical analysis is highly problematic.” Further, “On the whole the comparison between the outcome after prosthetic implantation and fusion as stated in this publication therefore cannot stand.”
- Some design details of the devices are stated incorrectly.
- The number of sizes of devices has increased, “which should reduce the incidence of subsidence.”
- The indications for lumbar disc arthroplasty have narrowed.
- Surgeons have gained more experience with surgical technique and implant placement which should lead to better outcomes.

The long term results of Lemaire (a minimum of 10 years) are noted by Dr. Buttner-Janz as being positive. One of the concluding remarks is, “Thousands of patients have been helped by the Charite artificial disc worldwide.” Reference was also made that if the logic of Putzier, et al., is used, organ transplantation and cardiac bypass surgery would not be options.

While Dr. Buttner-Janz offered explanation as to the noted outcomes, it does not change that Putzier is the longest follow-up to date (17 years) of the Charite disc, a device that was designed to last the life of the patient.

2. An attorney representing Stryker Corporation had these additional comments.

From the Charite FDA IDE trial:

“CMS attempts to discount these results, stating that at 24 months the Charite and fusion groups were equivalent with respect to VAS and ODI scores (Decision Memorandum, p.9). However, this eventual equivalence does not diminish the importance of the earlier positive results achieved with the Charite. The faster relief of symptoms with the Charite support the conclusion that it is reasonable and necessary as an option for treating DDD.”

CMS does not discount any valid results that may benefit the patient, however, the study was not designed to demonstrate faster symptom relief; this may or may not be true.

“Although CMS discounts satisfaction self-reporting as subjective and responsive to the placebo effect, the value of patient satisfaction should not be disregarded.”

CMS does not discount patient satisfaction, but rather the method of data collection and reporting in this study. This analysis technique causes uncertainty in interpretation.

“According to Blumenthal, et al., a higher proportion of fusion patients (85.9%) in the IDE study reported use of narcotics to control pain during follow-up, as compared to Charite patients (72.2%). Among patients demonstrating clinical success, there was a significantly lower rate of narcotic use for pain at 24 months in the Charite group (64%) than in the fusion group (80.4%).” “...Charite patients clearly benefit from the lower probability of requiring long-term medication usage as compared to fusion patients.”

CMS believes the high percentage of post operative narcotic use in both groups can not be construed as demonstrating benefit for the patient. Additionally, it may confound the results.

3. Stryker and DePuy identified these additional articles not previously mentioned as references:

Regan 2005 wrote about spine biomechanics, disc indications and contraindications, 2 case examples, preoperative planning, operating room preparation, operative technique with pictures, post-op period and rehabilitation, complications, some of the FDA IDE results, and also devoted a paragraph to the minimum 6 month follow-up of 100 consecutive patients, some of whom were in the FDA IDE trial. Little detail of the latter is provided.

Tropiano, Huang, et al., 2006 is not relevant to the focus of this analysis which is LADR with the Charite lumbar artificial disc.

SariAli, Lemaire, et al., 2005 reported on 17 patients (ages 31 to 42 at time of implantation) at a follow-up of 10.8 to 14.3 years post Charite implantation. The goal of this study was to develop a new technique in order to analyze axial rotation of a spinal segment, in vivo, after total disc replacement. Eleven patients had normal mobility (as compared to healthy controls) in torsion; whereas 6 had an abnormal increased mobility. No clinical correlates were given. This study does not provide additional evidence of benefit as its purpose was to develop new diagnostic techniques.

Denoziere and Ku 2006 compared, under severe loading condition, the biomechanics of the lumbar spine treated either by fusion or total disc replacement using computer simulation. The results were “The level implanted with the artificial disc showed excessive ligament tensions (greater than 500 newtons), high facet pressures (greater than 3 megapascals), and a high risk of instability. The mobility and the stresses in the level adjacent to the arthroplasty were also increased. In conclusion, the model for an implanted movable artificial disc illustrated complications common to spinal arthroplasty and showed greater risk of instability and further degeneration than predicted for the fused model.” This study again raises concerns about the long term complications of LADR.

Baldwin 2002 is a review of lumbar disc disease. Moore, Pinto, et al., 2002 and Hacker, Cauthen, et al., 2000 are articles on various fusion procedures.

Guyer and Ohnmeiss 2003 was a review article that compiled results and complications related to various types of lumbar and cervical disc replacements before the publication of the FDA IDE trial. They stated the rates of favorable results achieved were between 63% to 85%, with most reporting good results in 70% of patients. They reported a summary of complications from published literature, and stated, "Complications related to the use of the Charite III disc are similar to those encountered with anterior lumbar interbody fusion and are generally related to the approach to the disc space." He also noted, "As with reviewing complications encountered with any surgical procedure, the process is difficult because of the variation in the criteria each author uses to define a complication."

This article is illustrative of the difficulties in any spinal surgical procedure.

Anderson, Rouleau, et al., 2004 was a review article of cervical and lumbar arthroplasty. The authors concluded, "The early results are satisfactory, but the basic premise that motion preservation will diminish adjacent segment degeneration is yet unproven. Long-term results are unavailable and failure modes are unknown. Before implantation, the surgeon and patient must understand the experimental nature of the devices."

It is difficult to judge from this article what exact circumstances need to be met for a high probability of a successful outcome.

Guyer, McAfee, et al., 2004 was a report of 144 patients implanted with the Charite disc from 2 centers that were enrolled in the FDA IDE study, where the complete study report was given in Blumenthal, et al., 2005, McAfee, et al., 2005, and Geisler et al 2004.

This report does not provide any new evidence or insights not already taken into consideration in the subsequent study reports.

Walsh, Hanscom, et al., 2003 was an evaluation of the responsiveness of general and condition specific health status instruments, including the ODI and SF-36, for patients with low back pain/leg symptoms. Conclusions included that for studies of patients with low back problems, the general SF-36 may be a sufficient measure of health status and patient function, and pain scales appear to be the most responsive measure in patient with low back pain.

4. These additional comments are provided by surgeons and the general public:

Two references were provided for back surgery in general. In Deyo and Mirza 2006 (similar to Deyo, Gray et al. 2005), these issues were noted:

- Rates of spine surgery vary approximately fivefold among industrialized countries. The spine surgery rate in the U.S. is the highest in the world, and is approximately five times greater than the rate in England and Scotland.
- In the U.S., overall rates of spine surgery varied six fold among geographic regions whereas fusion rates varied 10-fold.
- The wide geographic variations in care suggest patients with similar characteristics may receive different procedures, depending on where they are and who they see.

Articles were referenced on lumbar fusion surgery and complications (Onesti 2004; Diwan, Parvartanei et al. 2003; Chen, Lai et al. 2003), results of procedures/approaches and/or revision (Pradhan, Nassar et al. 2002; Ondra and Marzouk 2003; Buttermann, Glazer et al. 1997; Etminan, Girardi et al. 2002; Christensen, Thomsen et al. 1998; Albert, Pinto et al. 2000; Slosar, Reynolds et al. 2000).

The above articles draw attention to the difficulties with spinal fusions in general. A review of this topic is outside the scope of this NCD.

A case report by David 2005 discussed revision of a Charite disc to a new Charite disc 9.5 years after implantation. The patient presented with moderate low-back pain and sciatica. A fragmented polyethylene core was found in a patient who was 52 years of age at the time of revision. The author concluded, "However, a revision necessitating an anterior approach carries significant risk to the vascular structures, the ureter, and neurological elements. It should only be performed by surgeons with a high degree of skill and experience in anterior lumbar surgery."

A 2004 abstract by Scott-Young was mentioned that reported on rate of revision for a series of 182 Charite patients since 1997 implanted by a single surgeon. Five cases required revision. Demographics and follow-up time are not indicated.

One commenter mentioned that ECRI had reviewed this technology in their TARGET database which monitors new developments on technologies. Some of their conclusions from the February 2006 update included:

"Limited data from these studies suggest that AIDR [artificial intervertebral disc replacement] may offer some advantages over spinal fusion and indicate that the short-term adverse-event rate for AIDR may be similar to that for spinal fusion. However, the true rate of complications and the clinical impact of complications cannot be determined yet because of the small numbers of patients studied, the various implants studies, the changing implant designs, and the limited data set. Furthermore, the available two-year safety data on AIDR are inadequate to allow conclusions to be drawn about long-term safety of AIDR compared to spinal fusion." These conclusions highlight the difficulties of the current literature.

There was a reference to a presentation, "Dr. Polly presented a meta-analysis of recent class I randomized controlled trials supporting the efficacy of surgical treatment of degenerative disc disease using both quality of life and cost parameters (North American Spine Society 2005)."

CMS was unable to locate a published study.

D. Comments without Evidence

The general issues identified in the comments are summarized below. Since almost 79% of the total comments received were a form letter, the percentage of commenters is significantly inflated for the first three issues because these issues were identified in the form letter.

CMS decision may influence other payers.

Approximately 82% of the commenters supported their opposition to the CMS non-coverage decision because it might influence other payers. While many of these commenters stated that this procedure was not applicable to most of the Medicare population, they expressed concern that other payers would follow the Medicare policy. One commenter stated, “It is unfair and would negatively impact patient care for CMS to look only at the Medicare population and formulate an adverse opinion that may affect patients of all ages and multiple payment sources.”

Our NCD process is based on a thorough review of the evidence and how that evidence relates to the Medicare population. Our analysis and decisions are available to the public. The conclusions and coverage determinations are for our Medicare population. Although CMS is aware that other payers may choose to follow Medicare coverage policy, this is not, and can not be, a consideration in the NCD process.

CMS decision may impact the future technology development and/ or coverage of new technology.

Concern was expressed in approximately 79% of the comments that the CMS non-coverage decision would negatively impact the future development and/or coverage of the artificial disc technology and technology in general. One commenter stated, "...the total disc technology is our future and should not be censured by a governmental agency."

CMS supports the development of new technology and evaluation of approaches that will improve health outcomes for the Medicare population.

Fusion is already covered. As an alternative to fusion surgery, the motion preservation aspect of LADR reduces or eliminates the degeneration of adjacent levels.

This was referenced in approximately 79% of the comments. Some suggested outcomes from fusion were poor so there should be an alternative. One commenter said, “The ability to regain motion at the operative level after surgery is a unique benefit of artificial disc replacement.”

We did not evaluate the long term complications of spinal fusion in this NCA, though we are aware of difficulties surrounding spinal fusion. This is a complex disease to treat, with some patients recalcitrant to current treatments. The use of fusion itself is variable with unclear indications and unreliable outcomes. However, among studies of the lumbar artificial disc, none have confirmed that patients experienced a noticeable benefit from what is referred to as regaining motion of the lumbar spine.

Based on the indications for use, the non-coverage decision may be appropriate for the 65 and over population but not for the under 65 population.

Approximately 4% of the commenters agreed that non-coverage may be appropriate for the 65 and over population; however, they felt that some of the Medicare population under 65 may be appropriate for this procedure. A number of commenters felt that patient selection based on strict indications for use was more appropriate than excluding the 65 and older population as is evidenced by this comment, “Age alone is not a contraindication to the implant and in fact de-facto restricts access to care.” Another commenter stated, “There is legitimate concern surrounding the paucity of data in the elderly population. Certainly, the role of total disc replacement in the elderly will remain relatively limited. ...CMS’ disabled population (over 6 million) represents an entirely different population. The results of the IDE study are directly applicable to this population.” DePuy spine commented, “The Level I data strongly suggests that there are beneficial outcomes for properly selected patients in the Medicare population and they should not be denied access to this technology.” One commenter stated, “...the procedure is not indicated in the elderly. In younger individuals on medicare, the vast number of patient’s I am familer [sic] with in this category have one or more contra-indications to disc replacement and thus I find that the denial of disc replacement surgery to the group of patients insured by medicare is not likely, in all medical probability, to represent even a minor constriction in surgeon decision making abilities nor a significant denial of patient care.”

CMS agrees there is a paucity of data for the 65 and older population and the FDA IDE Charite trial only included patients up to age 60, despite the fact that degenerative disc disease is most prevalent in those over 65.

The Decision is best left to the surgeon and the patient.

Some commenters felt that a non-coverage determination interfered with the doctor-patient relationship, equating noncoverage with CMS engaging in the practice of medicine, as was evidenced by this comment; “Any interference by the government into this decision-making process, not only interferes with the doctor-patient relationship, but is paternalistic. Who, but the patient has the right to decide whether the potential adverse events associated with a particular surgery outweigh the benefits?...CMS has no business engaging in the practice of medicine, which is essentially what they are doing by proposing a national non-coverage determination....The only way that surgeons, patients, and healthcare agencies in the United States are ever going to have availability of good quality scientific medical research regarding this device is if we generate it in the United States.” Some advocated that the decision should be left to the surgeon. One commenter felt that qualified physicians should be allowed to make the decision regarding treatment options for their patients and surgeons should be subject to an audit by a panel of their peers to review the indications. Many commenters felt that patient selection was key to good outcomes. However, one commenter stated, "Based on previous experience with new technologies, these devices are likely to be implanted in many individuals who should not be considered good candidates."

CMS only makes reasonable and necessary determinations under §1862(a)(1)(A) for payment purposes.

Promote coverage with restrictions.

A number of commenters supported coverage with restriction. DePuy has asked CMS to consider coverage with the following limitations.

The inclusion criteria provided by DePuy are consistent with FDA labeling, but differ from those of the PMA trial (age criteria 18-60). Symptomatic DDD confirmed by discography is included though this is of uncertain benefit (Cohen, Larkin et al. 2005). Exclusion criteria are similar to the PMA trial. In the DePuy criteria, DXA scans are required for those 50 and over. Under inclusion criteria, page 19, is listed, "Exclude patients with a score less than (1.0) SD below the norm, which indicates osteopenia. This would, by default, also exclude patients with a score less than (2.5) SD below the norm, which indicates osteoporosis"; and, under patient contraindications on page 23: "Osteopenia/Osteoporosis with a T-factor of < 1 ." This proposal is unclear for the following reasons:

- It is not clear if T-score or Z-score is intended. The two are not equivalent. A Z-score compares an individual with age-, gender-, and ethnicity-matched norms but is not used for diagnosis because a person's Z-score can remain constant throughout life, even as bone mineral density (BMD) declines with age. The T-score is the patient's BMD minus the young normal mean divided by the standard deviation of young normal mean. The World Health Organization developed a classification system for osteoporosis based on BMD using the known gradient of the risk of fracture in the population as a whole. These categories are based on T-score. The WHO classification system is derived from studies of white postmenopausal women and applies to them, but not for men, premenopausal women, or non-white postmenopausal women (Surgeon General's report 2004).

- The FDA IDE trial excluded persons with osteoporosis, osteopenia, and metabolic bone disease (Blumenthal et al 2005). It is not clear how osteoporosis or osteopenia is defined in this analysis, for in Regan 2005 (same trial), contraindications include: “Osteoporosis is defined as bone mineral density of more than 1 standard deviation below the norm for matched age group.” Note that this is not the standard definition for either osteoporosis or osteopenia.
- The assumption is that low BMD will predict poor arthroplasty outcomes though no evidence has been presented to support this premise. Dr. Geisler also suggested, “a DEXA scan to verify bone quality and eliminate osteopenic/osteoporotic patients.” BMD does not verify bone quality (NIH consensus report 2001, Surgeon General’s report 2004). BMD has been used to help predict fracture risk; it is compromised bone strength, the integration of bone density and bone quality that predisposes a person to an increased risk of fracture. There is currently no accurate measure of overall bone strength, so bone mineral density is used as a proxy, and accounts for about 70% of bone strength (NIH consensus report 2001). While BMD remains the single best predictor of fracture risk, “assessing the risk of bone disease and fracture remains a challenge (Surgeon General’s report 2004).” Further, “Not all of the risk factors have been identified, and the relative importance of those that are known remains unclear (Surgeon General’s report 2004).”

As discussed above, one of the goals of the NCD process is to clearly define those patient populations for which there is a clear health benefit. DePuy has proposed interesting criteria that may in fact more accurately identify appropriate patient populations, but this remains indeterminate unless tested through the appropriately designed clinical studies.

Promote coverage through continued collection of data.

A few advocated that CMS should provide coverage with evidence development.

Coverage with evidence development is a concept that was recently introduced and was intended to be used only in very limited circumstances where there was available evidence that was very promising that would, with some additional evidence collection, support a coverage determination for the Medicare population. This comment does illustrate that much remains unknown about this topic, and that more information is needed.

“The Charite device was significantly more effective than fusion in restoring the height of the collapsed disc space.”

CMS was unable to identify data that demonstrated that the patient experiences a noticeable health benefit from “restoring the height of the collapsed disc space”.

Complications are less with the disc as only 5.4% of Charite patients required a revision compared to 9.1% of the fusion group.

While Geisler noted equivalence in neurological complications between BAK and Charite, patients in the Charite group had a higher percent of severe or life threatening events (15% v. 9%) and device related adverse events (7.3% v. 4%), while device failures were higher in the BAK cage group (8.1% v. 4.9%). The quoted revision rate only looks at short term revision at the surgery index level, so it is possible that over time the revision rate for the Charite might surpass fusion, though this is currently unknown. The Charite complications in comparison to current fusion techniques is unknown.

In regards to the concerns about revision surgery, the primary revision approach is posterior, thus eliminating the vascular concerns of a repeat anterior approach.

Some lumbar artificial disc studies have noted that posterior revision may give unsatisfactory results. Additionally, studies have noted concern with a repeat anterior approach to remove or replace the prosthesis, which in some instances is unavoidable.

Ortho Device panel overwhelmingly approved the product

The FDA Ortho Device panel voted to approve the device but with certain conditions, such as a training requirement and post-approval studies, because the long-term safety and effectiveness of the device was unknown.

Reasonable and necessary is not defined.

While “reasonable and necessary” has not been defined in regulation, CMS has announced publicly and through these decision memoranda the framework we use to interpret “reasonable and necessary.” Appendix A in the proposed decision memorandum and in this document outlines the general approach which has been used for several years.

Insufficient weight was given to the FDA pre-market approval (PMA) study and excessive weight to less rigorous evaluations of the product including case series data and expert opinion.

DePuy spine and several commenters felt that the Blumenthal 2005 study Level I evidence was given insufficient weight and that the case series data, which were older, contained out of date information because technique and selection criteria were still being refined. DePuy Spine noted reliance on Level V evidence - expert opinion.

CMS relied primarily on the Charite IDE noninferiority study for the decision. Appendix A provides general methodological principles of study design and lists some of the methodological attributes associated with stronger evidence, including randomized controlled trials. It is further stated, "Methodological strength is, therefore, a multidimensional concept that relates to the design, implementation and analysis of a clinical study." Additionally, it is obvious that consideration must also be given to the results themselves, not simply the design of the study. CMS did ascribe sufficient weight to the Blumenthal study.

Other studies are mentioned primarily for complication rates and longitudinal trajectory of outcomes of patients. These are important considerations in a device that should last the life of the patient, and where there is significant question as to complication rates for a treatment of a non-life threatening disorder. Recent review articles (2005) continue to reference these earlier studies, as do other public comments. The level V information identified by DePuy includes a systematic review article in the evidence section and two references in the analysis section to authors of editorials, to give credit to the authors as is the common use of citations. It does point out, however, the paucity of high quality evidence, despite over 15,000 patients worldwide having the Charite (over 4,000 in the US alone).

Non-inferiority is an accepted statistical design for a clinical trial of an implant because the study compares the investigational device to the established standard of care, that is, safe and effective treatment for the condition in question.

The attorney representing Styker corporation further stated, “demonstrating non-inferiority versus an active control treatment should be sufficient to establish that the device is “reasonable and necessary.”

“A noninferiority trial seeks to determine whether a new treatment is no worse than a reference treatment” (Piaggio, Elbourne et al. 2006). Noninferiority trials present particular difficulties in design, conduct, analysis, and interpretation (Piaggio, Elbourne et al. 2006) and “Noninferiority and equivalence randomized trials create challenges for researchers and clinicians and are associated with several issues that are controversial and difficult to grasp, even for trialists” (Gotzsche 2006). One of the problems in the design of a noninferiority trial is illustrated by choice of control. The reference treatment’s efficacy should be established or in widespread use so that a placebo or untreated control group would be deemed unethical. It is controversial whether spinal fusion meets this criterion, as mentioned above. It is even more controversial that a procedure that is no longer in use such as the BAK would be the control. A recent study has found that noninferiority trials are poorly reported (Henaff, Giraudeau et al. 2006). For instance, one of the recommendations from Henaff is that trials report confidence intervals of treatment difference (or treatment ratio) and whether they are 1- or 2-sided. This was not done in the Charite IDE trial. In summary, CMS does not agree that merely performing a noninferiority trial automatically establishes that an item or service is reasonable and necessary.

An absence of conclusive evidence on a new technology should not be the basis of a non-coverage decision.

DePuy spine commented, “The absence of broad, conclusive evidence does not equate to the absence of clinical value. CMS should not make national non-coverage decisions unless it has definitive clinical information that a product or service is irrefutably not effective or that it causes patient harm.”

CMS disagrees with this premise. CMS is charged with determining whether or not an item or service is reasonable and necessary for the diagnosis or treatment of an illness or injury in the Medicare patients and non-coverage determinations may be made when the evidence is insufficient to make that determination.

DePuy Spine questions CMS cost data.

CMS does not take into consideration cost data in making NCDs. The information on cost data that appeared in the proposed decision memorandum was for information purposes only and has since been removed for this final document.

Allow local contractors to make the decision.

CMS believes that we have insufficient evidence for older adults and have made a national decision to non-cover LADR with the Charite lumbar artificial disc for the Medicare population over 60 years of age. While our proposed decision recommended noncoverage for all Medicare beneficiaries, after further analysis, we believe the evidence to be insufficient to arrive at a conclusion for the population 60 years of age and under and thus we will change our proposed decision and not alter current coverage at this time.

CMS should apply the same standard for coverage as the FDA uses in approving a new device.

DePuy Spine believes that CMS should follow “the same standard as that applied by the FDA in approving a new device,” of “whether the item in question is safe and effective.” It should be noted that FDA approval was completed without the benefit of the entirety of the evidence reviewed here. For this reason, a finding in favor of coverage without a thorough review of the currently available medical evidence would have been inappropriate.

As discussed above, the FDA “safe and effective” statutory standard is a marketing standard. CMS’ “reasonable and necessary” standard used in coverage determinations is also statutory and is used for reimbursement decisions. We believe that our decision is consistent with the statutory language found at §1862(a)(1)(A).

Other general issues mentioned.

Twenty seven commenters identified themselves or a family member as having had the procedure. All but one were positive about their results and commented that it gave them back their lives or explained the difference the procedure has made in their daily activities. Some of the statements made by these commenters included, "...it changed my life for the better", "My surgery was a complete success, and I was back at work full-time 3 weeks after my surgery", "My life after Charite has been a blessing for me and my family". One commenter who had the procedure stated, "As the physicians should well know and as has been pointed out in report after report covering the overview of the Artificial Disc Replacement (ADR) procedure, it is fraught with serious risks to the patients as well as post-operative problems, the latter of which I am personally experiencing. ...I am personally experiencing post-operative problems that are threatening the future of my current employment and which have already diminished my quality of life." None of these patients identified themselves as Medicare beneficiaries.

We also had 15 comments from people who were planning on having the procedure. These comments generally opposed our non-coverage decision because it might influence other payors and they would not have coverage for the procedure. Only one of these commenters identified themselves as a Social Security disability beneficiary.

A number of physicians provided anecdotal data about their patient's outcomes for this procedure. One physician indicated he was in the process of collecting outcomes data to submit for peer review and publication. The reported outcomes were mostly positive. One physician stated, "While most of the patients that I have performed this procedure on have done incredibly well; a few have not done quite as well as the others. Looking back, one may identify a common feature of patients that do well versus those who do less well. This kind of experience and improvement in patient care can only come if the surgeon is allowed to perform the procedure in the first place. As a final note on patient selection, it is the single most important determinant of the patient's long-term outcome." Only a few physicians stated that they had performed the procedure on a Medicare beneficiary.

Public comments from providers and patients describing their personal experiences are informative. It is encouraging that physicians continue to collect data and we would strongly urge them to do so in the context of a well-designed study.

VIII. CMS Analysis

National coverage determinations (NCDs) are determinations by the Secretary with respect to whether or not a particular item or service is covered nationally under Title XVIII of the Social Security Act §1869(f)(1)(B). In order to be covered by Medicare, an item or service must fall within one or more benefit categories contained within Part A or Part B, and must not be otherwise excluded from coverage. Moreover, with limited exceptions, the expenses incurred for items or services must be “reasonable and necessary for the diagnosis or treatment of illness or injury or to improve the functioning of a malformed body member.” § 1862(a)(1)(A). This section presents the agency’s evaluation of the evidence considered and conclusions reached for the assessment questions.

CMS focused on this general question:

Is the evidence sufficient to conclude that LADR with the Charite lumbar artificial disc will have health benefits for low back pain due to degenerative disc disease in the Medicare population?

Identifying the cause for chronic low back pain is challenging due to the complexity of the spine and the poorly understood neurophysiologic mechanisms of pain sensation (Haldeman 1999). For a variety of reasons, the damaged intervertebral disc (common in middle age and universal in old age) is judged to be the cause of chronic pain in many patients with low back pain (Huang, Sandhu 2004). A reliable test to determine the exact cause of low back symptoms has yet to be developed. Therefore, treatment of symptoms relies primarily on subjective evidence and clinical judgment. It must be remembered that the majority of patients with low back pain will have acceptable results without surgery. Nevertheless, in some patients, the pain is persistent and can eventually result in functional limitations. Spinal fusion surgery is offered to patients who do not respond to conservative treatments. No universally accepted guideline exists to assist the physician in patient management. In fact, spinal fusion for the indication of discogenic pain is controversial due in part to the conflicting results from clinical trials. Advocates of surgery over non-surgical care generally cite a single randomized controlled Swedish trial, while critics cite difficulties with that trial and the results of a more recent trial. Deyo, Nachemson, et al., 2004 stated, "Fundamental problems plague the study of spinal fusion, including the lack of definitive methods to confirm a solid fusion, a weak association between solid fusion and pain relief, and the placebo effect of surgery for pain relief." For patients, improvement in short term pain and function from fusion is unacceptably variable, and long term results remain controversial. In spite of the many difficulties and ambiguities, the use of spinal fusion surgery in the United States is rapidly increasing (Deyo, Nachemson 2004). The artificial lumbar disc has been developed as an alternative to fusion, with the premise that segmental mobility will improve outcomes, as has been the case for artificial hip and knee replacements.

The earliest evidence for benefits and risks with the Charite lumbar artificial disc comes from case series studies. In the 2003 systematic review of case series studies, de Kleuver, et al., reported that the proportion of patients classified as having "good" or "excellent" results varied in the studies from 50% to 81%. However, these numbers were difficult to interpret as there was no comparison group and no standardized method of reporting to compare study outcomes (de Kleuver, Oner et al. 2003). Various complications were observed in between 3-50% of patients. Patients with poor outcomes are candidates for fusion, with or without removal of the implant; however, posterolateral fusion without removing the disc may give unsatisfactory results, and the repeat anterior spine approach is challenging even for the most skilled surgeons. The authors all seemed to agree that patient selection is very important, yet the question of who is most appropriate for the device remains unanswered from these studies. The disc has been promoted as an alternative to fusion; yet, even in these short-term studies, spontaneous and surgical fusion occurred. The disc is required to function for many years, so important information about longer term benefits and risks, such as satisfaction, adjacent segment problems, and rate of re-operations, could not be determined from these earlier studies.

The recent Charite randomized, controlled trial was performed as part of the PMA Application to the FDA. In the PMA trial, designed to demonstrate noninferiority, the Charite lumbar artificial disc is found by the investigators to be no worse than the comparison device –the BAK cage with iliac crest bone graft. As it has happened, this fusion method has fallen out of favor with surgeons. Of important note for patient care, other techniques have shown a higher success rate and better ODI and VAS scores than either group in the Charite clinical study (Zindrick, Lorenz et al. 2005; Mirza 2005; Button, Gupta et al. 2005). This directly brings into consideration the design of this trial – that of noninferiority. “Noninferiority and equivalence randomized trials create challenges for researchers and clinicians and are associated with several issues that are controversial and difficult to grasp, even for trialists (Gotzsche 2006).” In other words, this type of trial design is complex and can give results that are difficult to interpret. “A noninferiority or equivalence trial requires that the reference treatment’s efficacy is established or is in widespread use so that a placebo or untreated control group would be deemed unethical” (Piaggio, Elbourne et al 2006).” The BAK cage is not currently in widespread use, and furthermore the use of fusion is highly variable, controversial, and has unclear indications. In the Charite PMA trial, the noninferiority design is inappropriate to demonstrate a health benefit.

The Charite PMA trial defined a successful outcome using a complex measure. Clinical success in the trial was defined by four criteria: (1) more than 25% improvement in the Oswestry disability score at 24 months after surgery, (2) no device failure, (3) no major complication, and (4) and no neurologic deterioration. This composite outcome is unconvincing as a demonstration of health benefit, particularly when the following points are also considered: 1) only 57% of disc replacement patients and 46% of BAK fusion patients met these four limited criteria; 2) in patients who were considered a clinical success at 24 months, 64% of the Charite group and 80.4% of the control were using narcotics; 3) at 24 months the change in VAS and ODI did not differ statistically from control; 4) the SF-36 PCS and MCS composite scores did not differ statistically from control; 5) there was no difference in operative time or blood loss between the two groups. The mean duration of hospitalization did differ (3.7 days versus 4.2 days for control) in these highly selected patients, but the discharge criteria were not standardized. A point was made by the investigators that patient satisfaction was greater in the disc group than the fusion group, yet those who entered the study were obviously willing to receive the new technology, and perhaps could have been disappointed with receiving the older technology (Zindrick, Lorenz et al. 2005). A summary of satisfaction of those who did not participate in the trial and subsequently had fusion, for response comparison purposes, was not presented; it would have made these results more convincing (Zindrick, Lorenz et al. 2005). Additionally, a sensitivity analysis was done with various imputations for patients that did not have complete follow-up data. The technique of data imputation by the sponsor was "last value carried forward." If data scenarios are examined to impute missing data, one finds that in a worst case scenario, where the missing data favored BAK fusion success and failure for Charite for the missing observations, noninferiority criteria were not met (FDA in-depth statistical review for expedited PMA 2004). An additional, very important point must be mentioned in the context of this PMA trial. As with other new products that go through the FDA review, only short-term results (24 month follow-up) are provided. All the patients, however, do not end their treatment at 24 months; they continue on with this implanted biomechanical device that must function successfully for many years. This trial does not provide data to address this concern.

In specific consideration of the Medicare population (who are elderly, disabled, or both), study exclusion criteria of the Charite randomized, controlled trial limit the generalizability of results. For instance, no one over age 60 was included in the study and patients with osteoporosis, osteopenia, and metabolic bone disease were excluded. A PMA trial participant (Regan 2005) listed contraindications to include: "Osteoporosis is defined as bone mineral density of more than 1 standard deviation below the norm for matched age group." Note that this is not the standard World Health Organization definition of either osteoporosis or osteopenia, and there are no official recommendations for the BMD standard deviation values that should be used to diagnose osteoporosis in men and premenopausal women (Surgeon General's Report 2004). Data were not provided on how many patients were screened to arrive at the 375 enrolled. Patients eligible for the Charite disc implantation, using strict criteria, may be narrowly focused. A study by Huang found that of 100 consecutive patients who had lumbar surgery in one spine surgeon's practice, 95% of patients had one or more contraindications to disc replacement, with the mean number of contraindications of 2.5 per patient (Huang, Sandhu 2004).

The short-term adverse event data from the Charite randomized, controlled trial are not easy to interpret. The publicly accessible FDA website does provide additional information. In the reporting of adverse events, some events are self-limited and resolve without incident, whereas other events may need additional services and may not easily resolve, creating significant morbidity. Patients may have more than one adverse event, a circumstance which is not differentiated in the reporting, so a straightforward calculation of rate may not have a clear meaning. Also in this study, there is a distinction between device related complications and approach related complications, which may not reflect the over-all risk for the procedure. It is valuable for patients to have an accurate perspective of risk.

The adverse effects noted in the analysis by the FDA are of potential concern. While a rate cannot be calculated without the total number implanted thus far, one can consider rates from other studies. Interestingly, most of these complications in the MAUDE database occurred within 2 months of operation. Bertagnoli noted:

“Most of the complications in total disc replacement procedures are iatrogenic; wrong indications, poor implantation technique, and improper positioning of the implant are the most likely causes. Isolated device-related complications are rare (e.g., subsidence, body fractures, polyethylene extrusion, and problems due to polyethylene wear). Due to stringently controlled inclusion groups, small study populations, and lack of long-term follow-up, only limited data are available. Lessons learned from hip and knee arthroplasty, however, suggest that the incidence of complications increases with duration of follow-up” (Bertagnoli, Zigler et al. 2005).

van Ooij raised interesting points in his case series report of complications with a mean follow-up of 53 months (van Ooij, Oner et al. 2003). He pointed out that while the disc prosthesis was often compared to hip and knee prosthesis, the multidimensional motion of the spinal segment is totally different than that of a hip or knee joint, so a comparison may have significant limitations. The normal intervertebral disc has a shock absorbing function, but very little has been written about this. In his series, they saw seven patients that had degeneration at levels other than the operated one, where it was not present before surgery. It is unclear if this is the result of the degenerative disease progressing, or, the result of stresses on the adjacent levels. van Ooij stated, “Many questions still exist concerning the biomechanics of a disc prosthesis.” Other important points he raised concerned the anterior surgical approach. The great vessels were mobilized for prosthesis insertion. The dimensions of the plate determined the extent of dissection. He acknowledged that this potentially could create bleeding and thromboembolic risks for the great vessels, which has been reported. Concern has been raised about the long term behavior of the biomaterials in the spine. Hallab, in his article on spinal implant debris-induced osteolysis, expressed a related concern, “With the introduction of modular artificial disc replacements and new materials for orthopedic spinal implants, the effects of implant debris on local and systemic tissues remains and will likely increase as a clinical concern ”(Hallab, Cunningham et al. 2003). If disc arthroplasty fails, there are three options: posterior fusion; revision replacement; and, anterior fusion (Kostuik 2004). As expressed by van Ooij, many agreed that revision through a repeat anterior lumbar approach can be very dangerous because of the adherence to the great vessels and the nerve plexus. While removal may not always be necessary in disc arthroplasty failure, some suggested posterior fusion may not give satisfactory results (van Ooij, Oner et al. 2003).

The 10 year follow-up study by Lemaire does give some evidence of long term viability (Lemaire, Carrier et al. 2005). It was unclear though, if follow-up was systematic or when the measures listed were recorded. The standard measures of VAS and ODI were not reported, and the Modified Stauffer Coventry scale score was not reported, but rather a percentage “relative gain”. Though complications were reported, it was not clear how many patients were free from complications at follow-up. There was no correlation of mobility with outcome. It was unclear how these case series patients differed from his 1997 fifty-one month case series study, as the mean age and range were very close, the mix of male and female differed (1997 [68 M, 37 F], 2005 [41 M, 59 F]), drawing into question who was included.

Putzier raised several important points in his 2005 article (Putzier, Funk et al. 2005). Of the 53 patients, 83% were either surgically fused or spontaneously ankylosed, with only 17% having near normal function of the spine. The authors did state that many causes for spontaneous fusions must be considered, including preservation of the anterior longitudinal ligament, which currently is believed to be a trigger for ossification. Secondly, they mentioned that for correct implantation of the prosthesis a complete removal of the remaining disc tissue must be accomplished, including decortication of the vertebral endplates. This process releases osteoinductive substances, which may also have caused this high rate of ossification. Additionally, they offered that “since all of the reported patients suffered preoperatively from moderate to severe DDD, a progression of these processes after surgery can be assumed.” An important point was that Charite I, II, and III are very different in design, yet there was no significant difference in the clinical or radiographic outcomes. One of the primary goals of the TDR is to prevent adjacent segment degeneration, yet the overall percentage judged radiographically (17%) was comparable to the results of follow-up studies after fusion surgery, which may follow from the high rate of spontaneous ankylosis. Lastly, those with preserved segmental mobility were statistically less satisfied than those with spontaneous ankylosis or surgical fusion. The authors hypothesized that while the anterior column is addressed by TDR, degenerative disorders of the posterior elements are not addressed.

Importantly, the major premise of spine segmental motion preservation is that adjacent level disease is increased by fusion surgery, and therefore that motion preservation will prevent this. However, there is no good evidence yet to support this premise, and there is some indication motion preservation with the Charite disc may not have an improved health benefit. As Hassett reported in his study on ageing and the nonoperated lumbar spine, there was radiographic evidence of progressing osteoarthritis of 3% to 4% per year, without symptom correlation (Hassett, Hart et al. 2003). This rate has also been quoted as the risk of adjacent level disease after fusion.

Lastly, some patients with poor results from disc arthroplasty will require revision surgery. Revision with posterior fusion appears to give less than satisfactory results. Anterior fusion with required anterior reentry is challenging, as reflected in a comment by McAfee, “even the most experienced vascular access surgeon has difficulty with the formidable revision through a repeat anterior lumbar procedure” (McAfee 2004). The issue of revision is of intrinsic importance: the device is intended for long term use; if revision surgery is necessary, treatment should not be worse than this non life-threatening disease.

Conclusions

Chronic back pain from degenerative disc disease is complex and can be difficult to treat. Current surgical treatment modalities are controversial. After thoroughly reviewing the existing data for LADR with the Charite lumbar artificial disc, important questions remain regarding patient selection, adverse events, and long term outcomes. The Charite PMA trial was limited to patient ages 18 to 60 years old, excluding the age group with the highest prevalence of degenerative disc disease. Due to the lack of evidence of benefit for those Medicare beneficiaries over the age of 60, CMS will noncover LADR with the Charite lumbar artificial disc in this population.

Some evidence does exist for patients 60 years of age and under, though the results of the Charite PMA noninferiority trial are unconvincing as an adequate demonstration of health benefit and do not provide a sufficient basis for a NCD at this time. The Charite studies without a comparison group make it difficult to draw clear conclusions on the benefit of treatment, though some individual patients with this complex, potentially disabling problem may benefit. This makes it difficult for CMS to arrive at an appropriate conclusion as to whether this device is reasonable and necessary for the 60 years of age and under Medicare beneficiary. In our proposed decision memorandum released on February 15, 2006, CMS proposed noncoverage for all Medicare beneficiaries. In consideration of the difficulty in arriving at a clear conclusion of the benefit of this technology for the 60 years of age and under population, along with the strong opinion from the public about the complicated nature of appropriate patient selection and the need for some limited coverage, we are changing our proposed decision and removing the national noncoverage for this segment of the Medicare population. Therefore, for Medicare beneficiaries 60 years of age and under, we will continue current coverage at this time.

There is clearly a tremendous need for additional research on the treatment of degenerative disc disease to include the technology addressed in the NCD—the lumbar artificial disc—and other surgical procedures to include spinal fusion. Therefore, CMS will also convene a Medicare Coverage Advisory Committee at the earliest possible time to address the issue of spinal surgery for degenerative disc disease. We urge the spinal surgery community to discuss the current limitations of the evidence for benefit and to outline the steps needed to develop better evidence.

CMS is aware that there are several other disc technologies in FDA investigational device exemption clinical trials in the United States. As previously stated, CMS is evaluating LADR with a focus on the Charite lumbar artificial disc in this analysis, since this was the only disc implant that had FDA approval at this time. However, we anticipate that when other lumbar spinal disc implants receive approval from the FDA that CMS will, by external request or internal direction, open this NCD for reconsideration with a thorough review of the evidence for each new disc implant. Since this NCD focuses on LADR with the Charite lumbar artificial disc, Medicare coverage under the investigational device exemption (IDE) for other lumbar artificial discs in eligible clinical trials is not impacted.

IX. Decision

The Centers for Medicare and Medicaid Services (CMS) has found that LADR with the Charite lumbar artificial disc is not reasonable and necessary for the Medicare population over sixty years of age. Therefore, we are issuing a national noncoverage determination for LADR with the Charite lumbar artificial disc for the Medicare population over sixty years of age. For Medicare beneficiaries sixty years of age and under, there is no national coverage determination, leaving such determinations to be made on a local basis.

[Appendices \[PDF, 214KB\]](#)

¹ BAK cage is a hollow metal cage implanted into the disc space of the spine, usually packed with bone, to stabilize the spine and allow fusion of the vertebrae.

[Back to Top](#)

Bibliography

Anderson P, Rouleau J. Intervertebral disc arthroplasty. Spine, 2004;29:2779-2786.

Bao QB, McCullen GM, Higham PA, et al. The artificial disc: theory, design and materials. Biomaterials 1996; 17 (12):1157-67.

Bertagnoli R, Kumar S. Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. *Eur Spine J* 2002;11:S131-S136.

Bertagnoli R, Yue, J, Fenk-Mayer A, et al. The treatment of disabling single-level lumbar discogenic low back pain with total disc arthroplasty utilizing the ProDisc prosthesis: a prospective study with 2 -year minimum follow-up. *Spine*. 2005a Oct 1; 30(19):2230-6.

Bertagnoli R, Zigler J, Karg A, et al. Complications and strategies for revision surgery in total disc replacement. *Orthop Clin N Am* 2005b;36:389-395.

Bertagnoli R, Karg A, Voigt S. Lumbar partial disc replacement. *Orthop Clin North Am* 2005c; 36 (3):341-7.

Bertagnoli R, Yue J, Nanieva R, et al. Lumbar total disc arthroplasty in patients older than 60 years of age: a prospective study of the ProDisc prosthesis with 2-year minimum follow-up period. *J Neurosurg Spine*, 2006a Feb; 4(2):85-90.

Bertagnoli R, Yue J, Fenk-Mayer, et al. Treatment of symptomatic adjacent-segment degeneration after lumbar fusion with total disc arthroplasty by using the ProDisc prosthesis: a prospective study with 2- year minimum follow up. J Neurosurg Spine 2006b Feb;4:91-97.

Blue Cross and Blue Shield technology assessment. Artificial vertebral disc replacement April 2005.

Blumenthal S, McAfee P, Guyer R, et al. A prospective, randomized, multicenter food and drug administration investigational device exemptions study of lumbar total disc replacement with the Charite artificial disc versus lumbar fusion. Part I: evaluation of clinical outcomes. Spine 2005;30:1565-1575.

Boden D, Davis D, Dina T, et al. Abnormal magnetic resonance scans of the lumbar spine in asymptomatic subjects. J Bone Joint Surg 1990;72-A:403-8.

Bone Health and Osteoporosis: A Report of the Surgeon General (2004)
<http://www.hhs.gov/surgeongeneral/library/bonehealth/content.html>

Button G, Gupta M, Barrett C, et al. Three- to six- year follow-up of stand-alone BAK cages implanted by a single surgeon. Spine J 2005;5:155-160.

Caspi I, Levindopf M, Nerubay J. Results of lumbar disk prosthesis after a follow-up period of 48 months. IMAJ 2003;5:9-11.

Chang BS, Brown PR, Sieber A, et al. Evaluation of the biological response of wear debris. Spine J 2004; 4 (6 Suppl):239S-44S.

Cinotti G, David T, Postacchini F. Results of disc prosthesis after a minimum follow-up period of 2 years. Spine 1996;21:995-1000.

Cohen SP, Larkin TM, Barna SA, et al. Lumbar discography: a comprehensive review of outcome studies, diagnostic accuracy, and principles. Reg Anesth Pain Med 2005; 30 (2):163-83.

Cunningham BW, Gordon JD, Dmitriev AE, et al. Biomechanical evaluation of total disc replacement arthroplasty: an in vitro human cadaveric model. Spine 2003; 28 (20):S110-7.

David T. Lumbar disc prosthesis: five year follow-up study on 147 patients with 163 SB Charite prosthesis. Eur Spine J 2002; 11(suppl 1):S18.

de Kleuver M, Oner F, Jacobs W. Total disc replacement for chronic low back pain: background and a systematic review of the literature. Eur Spine J 2003;12:108-116.

Delamarter RB, Bae HW, Pradhan BB. Clinical results of ProDisc-II lumbar total disc replacement: report from the United States clinical trial. Orthop Clin North Am 2005; 36 (3):301-13.

Deyo RA, Nachemson A, Miza SK. Spinal –Fusion Surgery – The Case for Restraint. N Engl J Med Feb 2004;350 (7):722-726.

Diwan AD, Parvataneni HK, Khan SN, et al. Current concepts in intervertebral disc restoration. Orthop Clin North Am 2000; 31 (3):453-64.

Durbhakula MM, Ghiselli G. Cervical total disc replacement, part I: rationale, biomechanics, and implant types. Orthop Clin North Am 2005; 36 (3):349-54.

Enker P, Steffee A, McMillin C, et al. Artificial disc replacement: preliminary report with a 3-year minimum follow-up. Spine 1993;18:1061–1070.

Errico TJ. Lumbar disc arthroplasty. Clinical orthopaedics and related research 2005;435:106-117.

FDA guidance document for the preparation of IDEs for spinal systems. January 13, 2000.

<http://www.fda.gov/cdrh/ode/87.pdf>. Accessed 1-6-06.

FDA In-depth Statistical review for expedited PMA (P040006) Charite artificial disc, DePuy Spine, Inc. (Dated Feb. 23, 2004) http://www.fda.gov/ohrms/dockets/ac/04/briefing/4049b1_04_Statistical%20Review%20Memo%20JCC.pdf . Accessed 12-05-2005.

FDA Center for Devices and Radiological Health, Orthopedics and Rehabilitation Devices Panel Meeting, Friday September 9, 2005. <http://www.fda.gov/ohrms/dockets/ac/05/transcripts/2005-4181t2.DOC>. Accessed 2-02-2006.

FDA memorandum to CMS. Analysis of adverse events reported to MAUDE for the Charite artificial disc prosthesis. November 23, 2005.

FDA PMA memorandum. Clinical review 2004.

http://www.fda.gov/ohrms/dockets/ac/04/briefing/4049b1_03_Clinical%20Review%20Memo%20MAY.htm. Accessed 12-05-2005.

FDA Summary of Safety and Effectiveness Data. 2004. <http://www.fda.gov/cdrh/PDF4/p040006b.pdf>. Accessed 12-05-2005.

Geisler F, Blumenthal S, Guyer R, et al. Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: results of a multicenter, prospective, randomized investigational device exemption study of Charite intervertebral disc. J Neurosurg (Spine 2) 2004;1:143-154.

Geisler F. Surgical technique of lumbar artificial disc replacement with the Charite Artificial Disc. Neurosurgery 2005; 56 (1 Suppl):46-57; discussion 46-57.

Gibson JNA, Waddell G. Surgery for degenerative lumbar spondylosis. The Cochrane Database of Systematic Reviews 2005, Issue 2. Art. No.: CD001352.pub2. DOI:10.1002/14651858.CD001352.pub2.

Goffin J, Van Calenbergh F, van Loon J, et al. Intermediate follow-up after treatment of degenerative disc disease with the Bryan Cervical Disc Prosthesis: single-level and bi-level. *Spine* 2003; 28 (24):2673-8.

Gotzsche P. Lessons From and Cautions About Noninferiority and Equivalence Randomized Trials. *JAMA* March 2006;295(10):1172-1174.

Griffith SL, Shelokov AP, Buttner-Janz K, et al. A multicenter retrospective study of the clinical results of the LINK SB Charite intervertebral prosthesis. The initial European experience. *Spine* 1994;19:1852-1849.

Gunzburg R, Mayer HM, Szpalski M, et al. Arthroplasty of the spine: the long quest for mobility. *Eur Spine J* 2002;11:S63-S64.

Hägg O, Fritzell P, Nordwall A: The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J*, 2003 12:12–20.

Haldeman S. Low back pain. Current physiologic concepts. *Neurol Clin* 1999;17:1-15.

Hallab N, Cunningham B, Jacobs J. Spinal implant debris-induced osteolysis. *Spine* 2003;28:S125-S138.

Hassett G, Hart DJ, Manek N, et al. Risk factors for progression of lumbar spine disc degeneration: The Chingford Study. *Arthrit Rheum* 2003;48:3112-7.

Henanff AL, Giraudeau B, Baron G, et al. Quality of Reporting of Noninferiority and Equivalence Randomized Trials. *JAMA* March 2006;295(10):1147-1151.

Hilibrand AS, Robbins M. Adjacent segment degeneration and adjacent segment disease: the consequences of spinal fusion? *Spine J* 2004; 4 (6 Suppl):190S-4S.

Hirsch C, Ingelmark BE, Miller M. The anatomical basis for low back pain. Studies on the presence of sensory nerve endings in ligamentous, capsular and intervertebral disc structures in the human lumbar spine. *Acta Orthop Scand* 1963; 33:1-17.

Huang R, Sandhu H. The current status of lumbar total disc replacement. *Orthop Clin N Am* 2004;35: 33-42.

Huang RC, Girardi FP, Cammisa FP, Jr., et al. The implications of constraint in lumbar total disc replacement. J Spinal Disord Tech 2003; 16 (4):412-7.

Kostuik J. Point of View. Spine 1996;32:1000.

Kostuik JP. Complications and surgical revision for failed disc arthroplasty. Spine J 2004;4:289S-291S.

Krismer M. Fusion of the lumbar spine. A consideration of the indications. J Bone Joint Surg (Br.) 2002;84:783-794.

Lee CK, Langrana NA. A review of spinal fusion for degenerative disc disease: need for alternative treatment approach of disc arthroplasty? Spine J 2004; 4 (6 Suppl):173S-6S.

LeHuec JC, Kiaer T, Friesem T, et al. Shock absorption in lumbar disc prosthesis: a preliminary mechanical study. J Spinal Disord Tech 2003; 16 (4):346-51.

Le Huec JC, Mathews H, Basso Y, et al. Clinical results of Maverick lumbar total disc replacement: two year prospective follow-up. *Orthop Clin N Am* 2005;36:315-322.

Lemaire J, Carrier H, Ali E, et al. Clinical and radiological outcomes with the Charite artificial disc. *J Spinal Disord Tech* 2005;18:353-359.

Lemaire J, Skalli W, Lavaste F, et al. Intervertebral disc prosthesis: results and prospects for the year 2000. *Clin Orthop* 1997;337:64-76.

Link HD. History, design and biomechanics of the LINK SB Charite artificial disc. *Eur Spine J* 2002; 11 Suppl 2:S98-S105.

Manek NJ, MacGregor AJ. Epidemiology of back disorders: prevalence, risk factors, and prognosis. *Curr Opin Rheumatol* 2005;17:134-140.

Marnay T. Lumbar disc replacement, 7-11 years results with Prodisc. *Eur Spine J* 2002; 11(suppl 1):S19.

Mayer HM, Korge A. Non-fusion technology in degenerative lumbar spinal disorders: facts, questions, challenges. Eur Spine J 2002a; 11 Suppl 2:S85-91.

Mayer HM, Wiechert K, Korge A, et al. Minimally invasive total disc replacement: surgical technique and preliminary clinical results. Eur Spine J 2002b; 11 Suppl 2:S124-30.

Mayer HM. Total lumbar disc replacement. J Bone Joint Surg Br 2005; 87 (8):1029-37.

McAfee P, Cunningham B, Holsapple G, et al. A prospective, randomized, multicenter food and drug administration investigational device exemptions study of lumbar total disc replacement with the Charite artificial disc versus lumbar fusion. Part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. Spine 2005;30:1576-1583.

McAfee PC. The indications for lumbar and cervical disc replacement. Spine J 2004;4:177S-181S.

McAfee PC. Comments on the van Ooij article. J Spinal Disord Tech 2005; 18 (1):116-7.

Mirza S. Point of view: commentary on the research reports that led to Food and Drug Administration approval of an artificial disc. Spine 2005;14:1561-64.

National Institute for Clinical Excellence. Interventional procedures consultation document - prosthetic intervertebral disc replacement. 2003. London. <http://www.nice.org.uk/page.aspx?o=87531>. Accessed 12-05-2005.

National Institute for Clinical Excellence. Overview of Prosthetic Intervertebral Disc Replacement, 2003. London. <http://www.nice.org.uk/pdf/ip/ip126overview.pdf>. Accessed 12-05-2005.

National Institutes of Health Consensus Conference. Osteoporosis Prevention, Diagnosis, and Therapy. JAMA 2001;285,6; Research Library pg. 785.

Oskouian RJ, Whitehill R, Samii A, et al. The future of spinal arthroplasty: a biomaterial perspective. Neurosurg Focus 2004; 17 (3):E2.

Piaggio G, Elbourne D, Altman D, et al. Reporting of Noninferiority and Equivalence Randomized Trials; An Extension of the CONSORT Statement. JAMA March 2006;295(10):1152-1160.

Polly DW, Jr. Adapting innovative motion-preserving technology to spinal surgical practice: what should we expect to happen? Spine 2003; 28 (20):S104-9.

Pope M, DeVocht, J. The clinical relevance of biomechanics. Neurol clinics North Am, 1999;17:17-41.

Putzier M, Funk J, Schneider S, et al. Charite total disc replacement – clinical and radiographical results after an average follow-up of 17 years. Eur Spine J 2005: published online October 28.

Regan J. Clinical results of Charite lumbar total disc replacement. Orthop Clin North Am 2005;36:323-340.

Santos E, Polly D, Mehbod A, et al. Disc arthroplasty: lessons learned from total joint arthroplasty. Spin JH 2004;4:182S-189S.

Shuff C, An HS. Artificial disc replacement: the new solution for discogenic low back pain? Am J Orthop 2005; 34 (1):8-12.

Singh K, Vaccaro A, Albert T. Assessing the potential impact of total disc arthroplasty on surgeon practice patterns in North America. Spine J 2004;4:195S-201S.

Sott AH, Harrison DJ. Increasing age does not affect good outcome after lumbar disc replacement. International Orthopedics (SICOT) 2000;24:50-53.

Trout AT, Kallmes DF, Kaufmann TJ. New Fractures after Vertebroplasty: Adjacent Fractures Occur Significantly Sooner. Am J Neuroradiol 2006;27:217-23.

U.S. Department of Health and Human Services. Vital and health statistics, Series 6: Cognition and survey management No.6. 1992. http://www.cdc.gov/nchs/data/series/sr_06/sr06_006.pdf. Accessed 11/07/2005.

U.S. Preventive Services Task Force. Guide to clinical preventive services, 2nd ed. Baltimore: Williams & Wilkins, 1996: 699.

van Ooij A, Oner F, Verbout A. Complications of artificial disc replacement; A report of 27 patients with the SB Charite disc. J Spinal Disorders & Techniques 2003;16:369-383.

Zeegers W, Bohnen L, Laaper M, et al. Artificial disc replacement with the modular type SB Charite III: 2-year results in 50 prospectively studied patients. Eur Spine J 1999;8:210-217.

Zigler JE. Clinical results with ProDisc: European experience and U.S. investigation device exemption study. Spine 2003; 28 (20):S163-6.

Zigler JE. Lumbar spine arthroplasty using the ProDisc II. Spine J 2004; 4 (6 Suppl):260S-7S.

Zindrick M, Lorenz M, Bunch W. Editorial response to Parts 1 and 2 of the FDA IDE study of lumbar total disc replacement with the Charite artificial disc vs. lumbar fusion. Spine 2005;30:E388-E390.

[Back to Top](#)